Institut d'études politiques de Paris

ÉCOLE DOCTORALE DE SCIENCES PO

Programme doctoral en économie

Département d'économie

Doctorat en sciences économiques

Essays on Innovation and Trade

Stefan Pauly

Thesis supervised by

Thomas Chaney, Professor of Economics, Sciences Po

defended on June 29, 2023

Jury

Davide Cantoni (*referee*)
Professor of Economics and Economic History, Ludwig-Maximilians-Universität München

Thomas Chaney (*thesis supervisor*)
Professor of Economics, Sciences Po

Isabelle Méjean (*president*)
Professor of Economics, Sciences Po

Claudia Steinwender (*referee*)
Professor of Economics, Ludwig-Maximilians-Universität München

Acknowledgements

I am enormously privileged to have been supervised by Thomas Chaney, whose creativity is so unique and whose passion for research is so contagious. Thank you for always taking my crazy ideas seriously and for letting me branch out into fields and topics that sometimes seemed so far apart. The curiousity that you nurtured kept me going all those years.

I also want to thank Johannes Boehm for his incredible kindness and generosity. It was extremely important for me to know I could always knock on your door and get great advice, about my papers but also about life. Thank you also to Golvine de Rochambeau for always being willing to discuss ideas and for connecting me with the right people to start my work in Myanmar. Besides the academic staff, the department would not function without all the wonderful administrative staff, I want to thank especially Cathy, Hadjila, Pilar and Sandrine for always being incredibly helpful.

I also want to thank Davide Cantoni, Isabelle Méjean and Claudia Steinwender for accepting to be part of my jury.

Next, I want to thank my amazing coauthors. Fernando, I will keep our days in Biarritz in fond memory, when we started working together on what is probably the most ambitious project of this thesis. Besides being a great coauthor, I value the friendship we have. Vladimir, I enjoyed walking through Paris with you, thinking about how tourism interacts with local amenities, or just drinking a beer. Thank you also to Çağatay and Beata, who were always available for brainstorming and helped me get to the finish line.

My colleagues and friends in Paris from the Combles, thank you for the years together! I am thinking of those who were there when I first arrived at Sciences Po: Alaïs, Arthur, Aseem, Camille, Charles, Elisa, Jan, Julia, Julien, Nicolò, Oliver and Pierre. Thank you to

my great cohort of PhD students: Daniel, Edgar, Pauline, Victor and Zydney. It was also great to meet those who joined the party a bit later: Gustave, Leonard, Moritz, Samuel, Sophia, and many, many more.

Words cannot express how thankful I am for all the wonderful people that I can call my friends. All the people I met throughout my life, starting with my childhood friends: Anna, Chantal, Christoph, Julius, Katha, Mo, Rick, Tessa and Tilmann. Then, the people I met during my carefree time in Berlin: Corinna, Daniel, Gerry, Kilian, Konsti, Monique, Morten, Tilman and others. Then, those from the time when I got a bit serious about economics in Toulouse: Esteban, Fabio, Julius, Luise, Matthieu and others. I want to thank my flatmates, Michelle and Fine, who were crucial to my sanity, especially in the second half of this journey, and Mareike, I am so glad we met!

Finally, I want to thank my family, my brother Martin, Mirka, and my parents, Renate and Jürgen, to whom I owe everything.

Contents

List of Figures

List of Tables

Introduction

Starting in the late 1980s, with the inception of the endogenous growth literature, economists have recognized the important role of technological progress for sustained economic growth. In a landmark paper Romer (1990) models technological progress as a function of deliberate human capital investment in a nonrival, partially excludable good. The nonrivalry of ideas, a key feature of Romer's model, implies that new ideas create positive externalities. Knowledge spillovers have also long been recognized as a key agglomeration externality (Marshal, 1920). Yet, despite of their perceived importance, knowledge spillovers are hard to pin down empirically.

Jaffe et al. (1993) made an important step in this direction by introducing the use of patent data to the economic literature. They use patent citations to study the geographic localization of knowledge spillovers. Building on this early work and the subsequent literature, the first two chapters of this thesis use patent data to closely track how inventors collaborate and build on each others' ideas. In particular, patent data contains precise information on where inventor teams are located, which pre-existing knowledge they cite and rely on, and which firms file the patent.

The first chapter of this thesis, co-authored with Fernando Stipanicic, investigates how an increase in human mobility can lead to more knowledge spillovers across locations. We then study how these knowledge spillovers translate into the creation of new ideas. This helps to identify key parameters of the production function of ideas that the early theoretical literature was not able to pin down due to a lack of detailed data. In addition, we show that initially less innovative cities benefit the most from an increase in spillovers, thereby causing convergence.

Rather than studying knowledge spillovers and its effects at a fairly aggregate level, a recent strand of literature describes how knowledge is produced inside the firm. With the

rise of multinational firms, production has become increasingly dispersed across countries inside the same firm. At the same time multinationals are responsible for a large share of knowledge creation. The second chapter, which is joint work with Çagatay Bircan and Beata Javorcik, looks at how frictions to communication can affect the collaboration between affiliate-based inventors with those located at the firm's headquarters. This is crucial since a large stock of firms' knowledge is held by HQ-based inventors and, as we show, inventors in foreign affiliates file patents of higher quality if they collaborate with HQ. While the first chapter captures knowledge spillovers across firms, the second chapter thus looks at the determinants of knowledge diffusion inside the firm.

Compared to knowledge spillovers and the exchange of ideas, the exchange of physical goods has long been recognized as a key driver of modern economic growth. A plethora of papers studies the interaction between bilateral frictions to trade, the exchange of goods and its effects on welfare and growth. While much of this literature focuses on more permanent frictions to trade such as tariffs or geography, few studies look at how frequent, unpredictable disruptions to markets affect how firms trade. This is especially relevant in developing countries where infrastructure can be unreliable and trade is often dominated by informal, less resilient firms. I study in my third chapter how firms in Myanmar deal with a repeated loss of access to markets in neighboring countries caused by ethnic conflicts along major roads. I find that firms partially compensate losses by selling domestically and delaying shipments. Firm size and connections to the military make firms more resilient.

Chapter 1: The Creation and Diffusion of Knowledge: Evidence from the Jet Age

In the first chapter, which is joint work with Fernando Stipanicic, we look at how the mobility of people affects the diffusion of ideas. We study a major technological breakthrough which made travel easier: In the 1950s the jet engine was introduced to civil aviation which resulted in faster, longer-range airplanes. At the time the flight network was heavily regulated, constraining airlines in reoptimizing their routes, allowing us to argue that the improvements in technology resulted in a plausibly exogenous shock to the existing network.

To study this historical period we digitize flight schedules of the major airlines operating at the time. This allows us to compute the exact travel time by plane before and after the arrival of jet planes. We find that travel time between US cities decreased by one third

from 1951 to 1966. We combine this with data on patents assigned to firms and inventor locations. This gives us the full picture of a firm's patenting at a particular time, including where the patented knowledge was produced and which prior art it was citing.

We find that the large increase in mobility resulted in an increased flow of ideas, especially for long distances. Patent citations between establishments at a distance of more than 2000km increase by 8% as a consequence of the reduction in travel time, corresponding to 38% of the overall increase in that distance interval.

The increase in knowledge diffusion further led to the creation of new knowledge. Using the elasticity of knowledge diffusion to travel time estimated in the first step, we construct a measure of knowledge access. We show that the rate of patenting increases, as a sector in a particular location gains access to new knowledge through the decline in travel times. Moreover, the results provide evidence that this effect was stronger for initially less innovative locations, implying innovation convergence across locations.

In summary, this historical episode allows us to establish a causal link between travel time and knowledge diffusion. Lowering frictions to the mobility of people increases the flow of ideas. We show that increased knowledge diffusion is crucial to new innovation, especially for inventors located outside of the biggest innovation hubs.

Chapter 2: Time after Time: Communication Costs and Inventor Collaboration in the Multinational Firm

In the second chapter, which is joint work with Çağatay Bircan and Beata Javorcik, we look at knowledge production inside the global firm. To study the global R&D activities of firms, we combine patent data with firm ownership links, allowing us to attribute each patent to its ultimate owner and its location of production through information on inventors' location of residence.

We first document that a large share of firms' R&D happens outside of the HQ country. For example, only around 61% of the most valuable patents filed by US firms are produced by purely US-based inventor teams. Focusing on patents that involve a foreign affiliate, we show that inventions are of higher quality if they are created in collaboration with HQ. This suggests that accessing HQ knowledge can make affiliate-based inventors more productive and underlines the importance of frictions to this knowledge diffusion.

We find that larger time zone differences between an affiliate and headquarters make collaborations less likely. A doubling in business hour overlap increases the probability of collaboration by 5 percentage points, even more so in experimental technology classes that likely require a high frequency of interactions. The effect remains relatively stable after controlling for physical distance. Taken together, this suggests that time zone differences constitute a friction to communication, which is distinct from travel cost.

Motivated by an increasing share of global collaborative patents and to provide causal evidence on the role of communication frictions, we study liberalization episodes of the telecom sector in the 1980s and 1990s. This deregulation led to a significant drop in international call prices. Gathering data on liberalization dates and bilateral call prices, we show that collaborations between affiliates and headquarters increase when telecom markets are liberalized in both countries and bilateral call prices fall. We then show that the effect of cheaper international calls was concentrated in location pairs with a high overlap in business hours and in experimental technologies, consistent with our cross-sectional results. Finally, we show that the effect is concentrated in less experienced inventors.

In summary, we show that time zone differences are a barrier to knowledge diffusion inside multinational firms. While improved communication technologies can facilitate cross-border inventor collaboration, time zones will likely continue to matter. We show that this is especially true for technologies that require a high frequency of communication and less experienced inventors.

Chapter 3: Traders in Trouble on the Road to Mandalay

Finally, in a third paper, I look at how uncertainty in trade frictions affects the exchange of goods. I combine transaction-level trade data from Myanmar with data on conflict. Myanmar has a long history of ethnic conflict which is concentrated in its border regions. At the same time, it relies heavily on exports to its neighbors (China and Thailand, in particular), some of which travels via the land route and thus through the conflict-affected areas. Treating the emergence of insurgencies in border regions as sudden, unexpected shocks to trade costs, I confirm anecdotal evidence that conflicts negatively affect the trade passing through.

Next, I document that firms' exports typically trade via one border city only. This lack of diversification makes firms vulnerable to disruptions caused by conflicts along a particular

route. Accordingly, local prices of export goods drop when important routes are disrupted, suggesting that firms cease to export and instead sell some of their produce on local markets. Trade drops more for perishable products that are risky to export when travel time to the border is uncertain. For durable products I observe that firms bundle shipments and increase the average shipment size, shielding them from losses.

Using the perishability of products to identify how different firms operate under uncertainty, I find that large firms are more likely to continue exporting. In addition, connections to the military, a powerful actor in Myanmar, are especially valuable during times of conflict and for firms that trade perishable products.

In summary, this project shows that the combination of a lack of diversification in export markets, combined with frequent disruptions, exerts a large toll on gains from exporting affecting firms and sectors heterogenously. While an end to Myanmar's ethnic conflicts is not in sight, reducing entry barriers to new export routes and a strong domestic market would significantly increase the resilience of firms.

References

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Marshal, A. (1920). Principles of economics. *London: Mcmillan*.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102.

Chapter 1

The Creation and Diffusion of Knowledge: Evidence from the Jet Age

This chapter is co-authored with Fernando Stipanicic (UC Berkeley).

Abstract

This paper provides new causal evidence of the impact of air travel time on the creation and diffusion of knowledge. We exploit the beginning of the *Jet Age* as a quasi-natural experiment. We digitize airlines' historical flight schedules and construct a novel data set of the flight network in the United States. Between 1951 and 1966, travel time between locations more than 2,000 km apart decreased on average by 41%. The reduction in travel time explains 33% of the increase in knowledge diffusion as measured by patent citations. The increase in knowledge diffusion further caused an increase in the creation of new knowledge. The results provide evidence that jet airplanes led to innovation convergence across locations and contributed to the shift in innovation activity towards the South and the West of the United States.

JEL Classification: O31, O33, R41, N72

Keywords: Knowledge Diffusion, Innovation, Travel Time, Airplanes, Jet Age

1. Introduction

"If I have seen further it is by standing on the shoulders of Giants."

– Isaac Newton (1675)[1]

"(...) if one man starts a new idea, it is taken up by others and combined with suggestions of their own; and thus becomes the source of further new ideas."

– Alfred Marshall (1890)[2]

In their famous quotations Isaac Newton and Alfred Marshall illustrate that access to knowledge is key for the creation of new knowledge. Understanding the process of creation of new knowledge is crucial as it has been characterized as one of the main causes of economic growth (Lucas (1993), Aghion and Howitt (1997) and Jones (2002)). Access to knowledge spurs the creation of new knowledge (Furman and Stern (2011), Acemoglu et al. (2016)). Physical proximity, by facilitating face to face interactions, is a key driver of the diffusion of knowledge and hence of access to knowledge (Storper and Venables (2004), Glaeser (2011)).

Providing evidence of the effect of access to knowledge on the creation of new knowledge is an empirical challenge. Agents who are highly productive in terms of creation of knowledge may endogenously sort towards locations with high access to knowledge, leading to reverse causality. Additionally, access to knowledge is correlated with other drivers of innovation as access to markets, resulting in a potential omitted variable bias due to confounding factors.

This paper provides new causal evidence on this question by exploiting as a quasi-natural experiment the beginning of the *Jet Age* in the United States. During the 1950s the introduction of jet engines into civil aviation led to a large reduction in travel time. We exploit changes in travel time to identify changes in knowledge diffusion, which are further translated into changes in access to knowledge. Then, we exploit changes in access to knowledge to study the impact on the creation of new knowledge. The results provide evidence that jet airplanes led to innovation convergence across locations and contributed

[1] Quoted from a letter of Isaac Newton to Robert Hooke, 1675. A digital copy of the letter can be found at: `https://digitallibrary.hsp.org/index.php/Detail/objects/9792`

[2] Quoted from Duranton and Puga (2004), page 2066.

to the shift in innovation activity towards the South and the West of the United States.

We start by constructing a new dataset of the flight network in the United States during the 1950s and 1960s. We digitize historical flight schedules of the major interstate airlines operating in the period and obtain the fastest route between every two airports in the network.[3] We document that between 1951 and 1966 travel time decreased on average by 29%, and the decrease is on average of 41% for airports located more than 2,000km apart.[4]

This nationwide shock was arguably exogenous as it happened in a strictly regulated environment. We decompose the change in travel time and find that 90% of the change is due to the improvement in aircrafts' speed, while 10% is due to a change in the flight routes. This is consistent with the fact that during this period the Civil Aeronautics Board (CAB) was imposing strong regulation in the interstate airline market. With the objective to promote a *stable* airline industry, the CAB determined ticket prices and restricted entry of airlines into new or existing routes.

Additionally, during the 1950s and 1960s airplanes were predominantly used to transport people and not goods. Hence, the change in travel time represented a shock to the mobility of people while not significantly affecting the shipment of goods.

To study knowledge creation and diffusion we use patent data. We follow Jaffe et al. (1993) and use patent citations as our observable measure of knowledge flow. We assemble one dataset with all corporate patents granted by the United States Patent and Trademark Office (USPTO) with filing year between 1949 and 1968, which includes for each patent: filing year, technology classification, location (Metropolitan Statistical Area, MSA) of the inventors when they applied for the patent, owner of the patent and citations to other patents which were granted by the USPTO.

We document three facts of patenting activity during our sample period. First, patent growth was stronger both in initially less innovative MSAs and in MSAs in the South and the West of the US. Second, over time multi-establishment firms expanded geographically and accounted for a larger share of patents. Third, the mass of citations shifted towards longer distances. Our results show that the decrease in travel time contributed to all three

[3]The 6 domestic airlines in our data accounted for 75% of total air passenger transport.

[4]New York and Boston are about 300km apart, while New York and San Francisco are located about 4,130 km apart. Between 1951 and 1966 we observe a reduction of travel time of 23% (13 minutes reduction) between New York and Boston, while the reduction is of 50% (5 hours 30 minutes reduction) between New York and San Francisco.

facts.

We do our analysis in three steps. In the first step, we estimate a gravity equation to obtain the elasticity of citations to travel time. We identify the elasticity exploiting only within establishment-pair across-time variation in citations and travel time. The estimated elasticity implies that citations increased on average 2.4% due to the decrease in travel time between 1951 and 1966. We find that the absolute value of the elasticity is increasing with the distance between the citing and cited establishments. At a distance of more than 2,000km, the change in travel time implies an increase in citations of 6.9%. This accounts for 32.7% of the observed increase in citations in this distance range.

In order to rule out the possibility that the opening of new routes or the timing of adoption of jets at the route level was driven by variables that also affected knowledge flows, we perform an instrumental variables estimation. We instrument the observed travel time with a fictitious travel time computed by fixing routes to the initial time period and assuming in each year all routes are operated with the year's average airplane. Hence, changes in fictitious travel time are only due to the nationwide roll out of jets and is thus independent of decisions at the route level. The results do not change significantly, reflecting the reduced scope for endogeneity of travel time. In addition, the results are robust to controlling for potential confounding factors such as changes in highway travel time, telephone connectivity and flight ticket prices. Finally, the results also remain after restricting the sample to contain only establishments that existed in the initial time period.

In the second step, using the estimated elasticity of diffusion of knowledge, we compute a measure of knowledge access that is specific to each location-technology. The measure captures changes in knowledge access that are only consequence of the change in travel time. We use knowledge access as an externality that affects the production of new patents and estimate the elasticity of new patents to knowledge access. We identify the elasticity at the establishment level comparing only across time variation in patents and knowledge access across establishments within a location, conditional on aggregate technological trends. Thus, the identification is independent of location specific changes in local population or R&D subsidies. The estimated elasticity implies that the amount of new patents filed increased at a yearly growth rate of 3.5% due to the increase in knowledge access, which accounts for 79.5% of the observed yearly growth rate.

Given the reduction in travel time was larger for longer distances, the increase in knowledge access was stronger in locations geographically far from the initial innovation centers

located in the Midwest and the Northeast. Hence, by increasing access to knowledge, the reduction in travel time led to a shift in the distribution of innovative activity towards the South and the West of the US. The South and the West had an average yearly growth rate of patenting 2.1 percentage points higher than the Northeast and the Midwest during our sample period. The change in travel time explains 35% of the observed differential growth.

We find that the value of the elasticity of patents to knowledge access is bigger in magnitude for establishments located in initially less innovative locations. Within each technology class, we rank locations according to the amount of patents in the initial time period and split them into four quartiles. We find that the increase in knowledge access predicts a 4.5% yearly growth rate of patenting in locations in the lowest quartile of initial innovativeness, while it predicts a 3.4% yearly growth rate in the highest quartile. The difference in growth rates indicates that the increase in knowledge access acted as a convergence force between locations, and it can explain 21% of the convergence observed in the data. Results go in the same direction if we rank locations in terms of patents per capita.

Our results are robust to controlling for changes in market access by highway, changes in market access by airplanes and time changing telephone connectivity. Results do not change if we compute knowledge access using only knowledge located at long distances. Additionally, we present suggestive evidence that the results are not driven by a decrease in financial frictions.

In the third step, we uncover the sources of the increase in patenting. We find that most of the effect of knowledge access on new patents happens through two entry margins: entry of establishments of new firms and entry of subsidiaries of firms that expanded from other locations. The two entry margins are stronger in initially less innovative locations, meaning that convergence comes both from new firms and the geographic expansion of multi-establishment firms.

To more directly test the firm expansion channel, we study if a firm's subsidiary's location decision depends on travel time to headquarters. We estimate a probability model to analyze if the locations in which firms have inventors applying for patents depends on travel time to the firm's headquarters. We identify the change in the probability only from changes in travel time and locations in which the firm starts patenting or stops patenting. We find that the probability that a firm has inventors applying for patents in a certain location goes up when then travel time from that location to the firm's headquarters reduces. In addition, the change in the probability is stronger for potential recipient

locations that were initially less innovative, again highlighting the importance of this channel for convergence.

This paper contributes to multiple branches of literature. First, it contributes to the literature on agglomeration and knowledge spillovers. Agglomeration forces are usually understood as happening in a geographically localized manner (Glaeser (2011), Arzaghi and Henderson (2008)). The literature on tech clusters also documents this fact (Duranton et al. (2009), Kerr and Robert-Nicoud (2020), Moretti (2021)). The seminal paper Jaffe et al. (1993) finds that patent citations decay rapidly with distance. Our results show that jet airplanes allowed long distance knowledge spillovers, facilitating the development of tech clusters in other regions. The literature that provides evidence of knowledge spillovers usually focuses on changes in the supply of knowledge (Bloom et al. (2013), Acemoglu et al. (2016)). In our case we fix the supply of knowledge and focus on changes in the degree of accessibility.

We contribute to the literature on transportation by studying a new quasi-natural experiment that isolates a shock to the mobility of people. To do so we construct a new dataset that could be used to answer many other questions.[5] Other papers have studied the impact of transportation improvements on innovation. Agrawal et al. (2017) study the impact on innovation of a region's stock of highways, while Perlman (2016) uses 19th century data on locations' density of railroads. Andersson et al. (2017) and Tsiachtsiras (2021) do so using the historical railroad expansion in Sweden and France. Relative to them, we contribute by exploiting a quasi-natural experiment that allows us to isolate a channel of face to face interaction, with little scope for a trade channel. In contemporaneous work Bai et al. (2021) estimate the elasticity of patent citations to air travel time using the introduction of new airline routes in a more recent period, post deregulation of the airline market. Relative to them, we contribute by exploiting a set up in which the risk for endogeneity of travel time is limited. Our work is related to other literature which found that business travel affects innovation (Hovhannisyan and Keller (2015)), trade (Söderlund (2020)) and industrial activity (Coscia et al. (2020)). Also, air travel shapes collaboration between researchers (Catalini et al. (2020)).

The impact of transportation improvements in economic outcomes has long been a subject of study (Fogel (1963), Baum-Snow (2007), Michaels (2008), Donaldson and Hornbeck (2016), Jaworski and Kitchens (2019) and Herzog (2021)). Our convergence result contrasts

[5] Our dataset also includes international flights. We are currently digitizing more airlines to increase coverage both inside the US and internationally.

with previous studies on improvements in other means of transport. Pascali (2017) finds that the introduction of steam engine vessels in the second half of the 19th century led to an increase in international trade which contributed to economic divergence between countries. Faber (2014) finds that the expansion of the highway system in China led to a reduction of GDP growth of peripheral counties, with evidence suggesting a trade channel. While both papers emphasize a trade channel, in our set up the trade channel would not be of first order. Hence, we uncover a new effect of improved connectivity.

Finally, we contribute to the literature on firm's location decision. Our result about firms deciding their establishments' locations based on travel time to headquarters is comparable to the one found by Giroud (2013), who finds that a reduction in air travel time to headquarters increases plant level investment and total factor productivity. Similarly, Campante and Yanagizawa-Drott (2017) finds that firms' cross country investment decision depends on connectivity to headquarters.

The paper is structured as follows. First, we present a simple theoretical framework which lays the foundations of how to think about the creation and diffusion of knowledge. The framework shows the two key parameters to estimate. Second, we describe the historical context in which jet airplanes were introduced. Third, we present the two datasets that we use: travel times and patents. Fourth, we perform the analysis to estimate the impact of travel time on the diffusion of knowledge, the creation of knowledge, and firm's location decision. Fifth, we conclude.

2. Conceptual framework

This section lays out a simple theoretical framework to think about the creation of knowledge. The framework clearly shows the two key parameters to estimate empirically: the elasticity of knowledge diffusion to travel time and the elasticity of knowledge creation to knowledge access.

Following Carlino and Kerr (2015) we consider a production function of knowledge which includes external returns in the form of knowledge spillovers. Knowledge output of a firm depends not only on firm's specific characteristics as its idiosyncratic productivity and input decisions, but also on an externality due to knowledge spillovers. We consider a

production function of knowledge of the following form:

$$\text{New Knowledge}_{Fi} = f(z_{Fi}, inputs_{Fi}) \times \text{Knowledge Access}_i^{\rho} \tag{1.1}$$

where New Knowledge_{Fi} is the knowledge created by firm F located in i. The output of Fi depends on an *internal* component and on an *external* component. The *internal* component is the firm's idiosyncratic productivity z_{Fi} and choice of inputs $inputs_{Fi}$. The *external* component, Knowledge Access_i, represents knowledge spillovers to which all firms F in location i are exposed to. The degree to which this externality affects the production of knowledge is governed by the parameter ρ. If ρ is zero then knowledge spillovers have no effect on the creation of new knowledge. On the other hand, a positive ρ implies that, keeping productivity and inputs constant, an increase in the level of knowledge spillovers leads to an increase in firm F's creation of new knowledge.

A long standing literature studies the importance of knowledge spillovers for the creation of new knowledge.[6] The concept of knowledge spillovers goes back at least to Marshall (1890) who explains it as one of the agglomeration forces. Krugman (1991) refers to knowledge spillovers as one of the justifications for external increasing returns, and that the degree of spillovers are dependent on physical distance. The geographic decay of spillovers is grounded in the fact that not all knowledge is easy to codify, usually referred to as *tacit knowledge*, and geographic proximity increases the degree of knowledge spillovers by facilitating face to face interactions (Storper and Venables (2004), Glaeser (2011)). Hence, we consider the total amount of knowledge spillovers to which the firm F in location i is exposed to has the following functional form:

$$\text{Knowledge Access}_i = \sum_j \text{Knowledge stock}_j \times \text{distance}_{ij}^{\beta} \tag{1.2}$$

where Knowledge stock_j is the total amount of knowledge in location j (which is non-negative) that could potentially spill over to location i and distance_{ij} is a measure of distance from j to i. The amount of knowledge that spills over from j to i depends on distance and the degree with which distance impedes spillovers, governed by the parameter β. If β is zero, then distance does not affect knowledge spillovers from j to i and all locations perfectly share the same level of *Knowledge Access*. On the contrary, a negative β implies a decay in knowledge spillovers when distance increases. In other words, a

[6]The chapters of Audretsch and Feldman (2004) and Carlino and Kerr (2015) in the Handbook of Regional and Urban Economics provide an excellent review on the literature on knowledge spillovers, their geographic decay and how they affect the creation of knowledge.

negative β implies that if we reduce the distance from j to i while keeping every other distance constant, the amount of spillovers from j to i will weakly increase.

This theoretical framework bears resemblance to the concept of *Market Access* presented in Donaldson and Hornbeck (2016) and Redding and Venables (2004). If we interpret *Knowledge Access* as one of the inputs in the production function of knowledge, then Knowledge Access$_i$ could be interpreted as a measure of *Input Market Access*. This measure captures how cheaply firms in location i can access pre-existing knowledge, where the cost of accessing knowledge depends on distance between i and j. Also, *Knowledge Access* is similar to a measure of network centrality. The centrality of each location i (node) is the weighted sum of distance (edges) to every location, where the weight of each location is given by its knowledge stock.

The theoretical framework highlights the two parameters to estimate: ρ and β. Empirically, we use travel time as a measure of distance to first estimate β and then conditional on β we estimate ρ. Changes in travel time due to improvements in commercial aviation allow us to estimate both parameters. First, we use citations between patents as a proxy for the diffusion of knowledge. We estimate β by relating changes in travel time between research establishments to changes in citations between them. Second, we use the stock of patents filed by inventors in each location as proxy for each location's stock of knowledge. We construct a measure of knowledge access using the patent stock, travel times and the value of β. New patents in each location proxy for new knowledge. Changes in travel time lead to changes in knowledge access which allow us to estimate ρ.

3. Historical context

3.1. Air transport: jet arrival

The jet aircraft was first invented in 1939 for military use, with the German Heinkel He 178 being the first jet aircraft to fly. The first commercial flight by a jet aircraft was in 1952 by the British Overseas Airways Corporation (BOAC) from London, UK to Johannesburg, South Africa with a Havilland Comet 1. Nonetheless, given the amount of accidents of the Havilland Comet 1 due to metal fatigue, jet commercial aviation did not truly take off until the Boeing 707 entered commercial service in late 1958. The 24th of January of 1959 represented a major milestone in the jet era: American Airlines Flight 2 flew with a Boeing 707 jet aircraft from Los Angeles to New York, the first non-stop transcontinental

commercial jet flight.[7]

In 1951 New York City and Los Angeles were connected with a one-stop flight in 10 hours and 20 minutes. The flight had a forced stop in Chicago and was operated with the propeller aircraft Douglas DC-6, which had a cruise of 500 kmh. By 1956, New York City and Los Angeles were connected with a non-stop flight in 8 hours and 30 minutes. This was accomplished due to the introduction of the propeller aircraft Douglas DC-7 which had a cruise speed of 550kmh, and a change in regulation which increased maximum flight time of a crew from 8 to 10 hours within a 24-hour window.[8] In 1961, the route was covered with the jet aircraft Boeing 707 in a non-stop flight in 5 hours 15 minutes, reaching 5 hours 10 minutes in 1966. The Boeing 707 had a cruise speed of 1000kmh, cutting travel time from New York City to Los Angeles in half between 1951 and 1966.

3.2. *Air transport: moving people, not goods*

During the 1950s and 1960s, air transportation served to transport people but not goods. Figures 1.1 and 1.2 are images (edited for better readability) from annual reports of the Interstate Commerce Commission of 1967 and 1965 respectively. Figure 1.1 displays the amount of passenger-miles[9] for Air, Motor and Rail transportation from 1949 to 1966. We observe that, while transport of people by rail decreased and by motor remained relatively constant, transport of people by air multiplied by 6 in a 16-year period, which translates to around 12% compound annual growth. In 1966, air transport accounted for more passenger-miles than both rail and motor transportation together, reflecting the growing importance of this mean of transport.

Figure 1.2 shows shipments in ton-miles for the period 1939 to 1964 by different means

[7] The reader passionate of aviation history would enjoy reading the following New York Times article which tells the experience of the first transcontinental jet flight: `https://www.nytimes.com/2009/01/26/nyregion/26american.html`

[8] AA and TWA had transcontinental non-stop propeller flights scheduled since at least 1954. These flights were scheduled to take 7 hours 55 minutes, just under the maximum flight time allowed by regulation in domestic flights: regulation impeded pilots from being on duty more than 8 hours within a 24 hours window. Nonetheless, the propeller aircrafts used in these flights, the Douglas DC-7 and the Lockheed Super Constellation, overheated their engines due to excessive demand to cover the route in less than 8 hours. AA fought intensely until the CAB approved a waiver that allowed non-stop transcontinental flights to take up to 10 hours to accomplish the non-stop transcontinental flight. See page 16 of the edition of the 21st of June 1954 of the Aviation Week magazine `https://archive.org/details/Aviation_Week_1954-06-21/page/n7/mode/2up`

[9] Passenger-miles is a standard unit of measurement in transport, where one passenger-mile accounts for one person traveling one mile. The reasoning is the same for ton-miles, with one ton of goods traveling one mile.

of transport. Interestingly, we observe that air transport of goods, even if it increased, it accounted for less than 0.1% of transport of goods in 1964.[10]

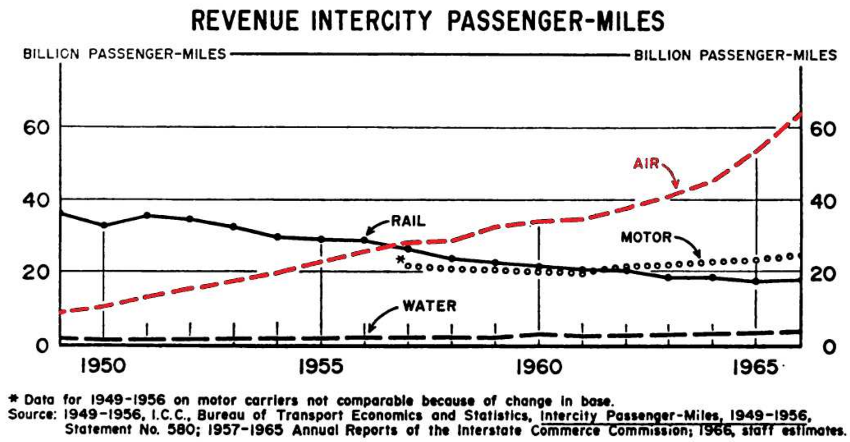


Fig. 1.1. Passenger Miles By Means of Transport

Source: Interstate Commerce Commission, Annual Report 1967. Edited by the authors.

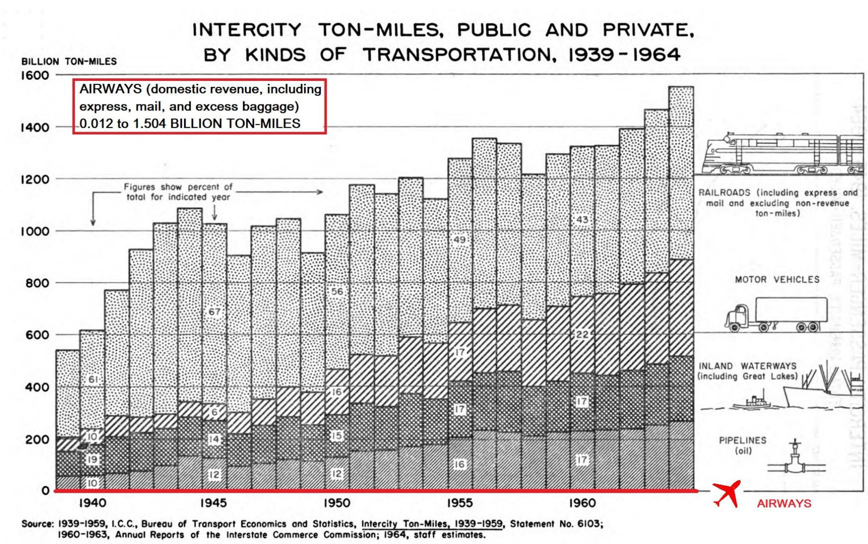


Fig. 1.2. Ton Miles by Means of Transport

Source: Interstate Commerce Commission, Annual Report 1965. Edited by the authors.

[10]We have not found data about shipments by mean of transport measured in monetary values.

3.3. *Regulation*

As explained in Borenstein and Rose (2014), in the 1930s the airline industry was seen as suffering from coordination issues, destructive competition and entry. Additionally, the industry was developing in a context of financial instability and increasing military concerns post Great Depression. A strong domestic airline industry was perceived as an interest of national defense. As consequence, the Civil Aeronautics Board (CAB) was created in 1938 with the objective to promote, encourage and develop civil aeronautics.[11] It was empowered to control entry, fares, subsidies and mergers.[12] In other words, the CAB regulated the market by deciding which airlines could fly, in which routes they could operate, the price that they charged in each route, the structure of subsidies and merger decisions. The CAB regulated the airline industry in a barely unchanged manner until it ceased to exist in 1985.

When the CAB was created, it conceived special rights to the existing airlines over the connections they were operating. The CAB did not permit entry of new airlines on interstate routes and gradually allowed current airlines to expand their routes. The CAB controlled both the system and each airline's network. The network was designed to maintain industry stability and minimize subsidies, leading to a system where each route was mainly operated by one or two airlines.[13] Importantly, Borenstein and Rose (2014) in pages 68-69 explain that *"the regulatory route award process largely prevented airlines from reoptimizing their networks to reduce operation costs or improve service as technology and travel patterns changed."* As a consequence, any technological improvement such as increases in aircraft speed, capacity or range would not affect each airline's flight network in the short term.

By regulating fares, the CAB explicitly encouraged airlines to adopt new aircraft. Airlines, when operating an older aircraft, would apply for a fare reduction arguing that it is needed in order to preserve demand for low quality service. The CAB would refuse this application, hence airlines would have to adopt new aircraft or risk losing consumers who would choose another airline which flies newer aircrafts.

[11] The CAB was a federal agency hence, in principle, would not have control over intrastate routes. Nonetheless, according to Borenstein and Rose (2014) the CAB managed to have some intrastate markets under its control using legal arguments.

[12] Safety regulation was under the control of the Federal Aviation Administration.

[13] Borenstein and Rose (2014) in page 68, based on Caves (1962), expose *"In 1958, for example, twenty-three of the hundred largest city-pair markets were effectively monopolies; another fifty-seven were effectively duopolies; and in only two did the three largest carriers have less than a 90 percent share."*

4. Air travel data

We construct a new data set of the flight network in the United States during the 1950s and 1960s. We collected and digitized information of all the flights operated by the main airlines and obtained the fastest route and travel time between every two airports in the network.

To construct the flight network we use historical flight schedules of the main airlines operating in 1950s and 1960s. Figure 1.3 is a fragment from an example page of the 1961 flight schedule of American Airlines. In the flight schedule we observe in the center column the name of departure and arrival cities (which we match to airports using airlines' historical ticket office geographical location), while the small columns on the sides depict flights. In the top of the small columns we observe the type of service provided (first class, coach or both), aircraft operated, days operated (daily if information is missing) and flight number. The content of the small columns displays the departure and arrival time (local time, bold numbers represent PM) at each city, including all intermediate stops.

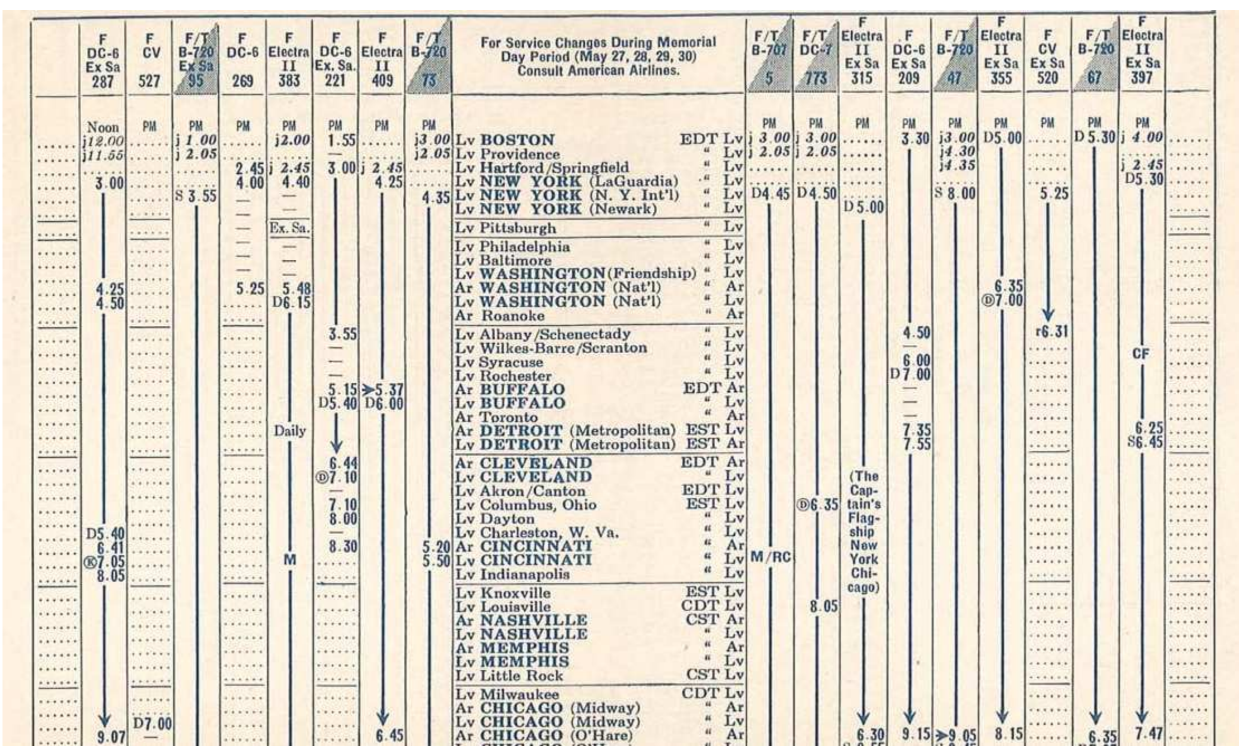


Fig. 1.3. Fragment of flight schedule American Airlines 1961

Source: American Airlines 1961. The center column displays the name of departure and arrival cities. The small columns on the sides display flights with departure and arrival time (local time, bold numbers represent PM). The top of the small columns shows the type of service provided (first class, coach or both), aircraft operated, days operated (daily if information is missing) and flight number.

We digitize flight schedules for the years 1951, 1956, 1961 and 1966 of six domestic airlines: American Airlines (AA), Eastern Airlines (EA), United Airlines (UA), Trans World Airlines (TWA), Braniff International Airways (BN), Northwest Airlines (NW),[14] and one international airline: Pan American Airways (PA). This group of airlines includes the *Big 4*: AA, EA, UA and TWA, which accounted for between 69% and 74% of interstate air revenue passenger miles in the US in the years collected. BN and NW were digitized in order to have a wide geographical coverage, while PA provides international flights. Based on C.A.B. (1966), in the years collected, the six domestic airlines together accounted for between 77% and 81% of interstate air revenue passenger miles.

In total we have digitized 6,143 US flights (unique combinations of flight number-year, 7,007 worldwide). However, flights often have multiple stops. If we count each non-stop part (*leg*) of these flights separately, our sample contains 17,737 legs in the US and 21,210 worldwide. Our data connects 275 US airports (434 worldwide) creating 2,563 unique origin-destination (directional) airport links (3,466 worldwide). Figure 1.4 displays the flight network in continental United States pooling all years together. In Appendix A.2 we show the US flight network by year, around 80% of the non-stop flights remain year-on-year.

Using departure and arrival time of each flight at each airport, we obtain the fastest route and corresponding travel time between every two airports in our data. To obtain the fastest route and travel time we modify the Dijkstra algorithm to account for layover time in case the fastest route includes connecting flights.

Once the fastest route between every two airports is computed, we match every airport to 1950 Metropolitan Statistical Areas (MSA) using the shape file from Manson et al. (2020). We consider only MSAs in contiguous United States. We use MSAs as the geographical unit of analysis because they are constructed taking into account commuting flows. We assume that people in an MSA would use, for each desired route, the most appropriate airport lying inside or nearby the MSA. We match each airport to all MSAs for which it lies

[14]These are six of the fifteen trunk (interstate) airlines operating in 1951. Many of the remaining trunk airlines would merge with another trunk airline over the years, and there would be zero entry of new airlines. We are currently digitizing the remaining trunk airlines and we plan to add them to the travel time dataset in the future. We have already digitized: Allegheny Airlines, Capital Airlines, Colonial Airlines, Continental Airlines and Delta Air Lines. We have also digitized the year 1970 for the six airlines used in this paper and Pan American. Due to a time constraint we have not included them in the current analysis. We plan to digitize BOAC to obtain more international flights, and to cover the time period 1930 to 2000 for all airlines that is possible.

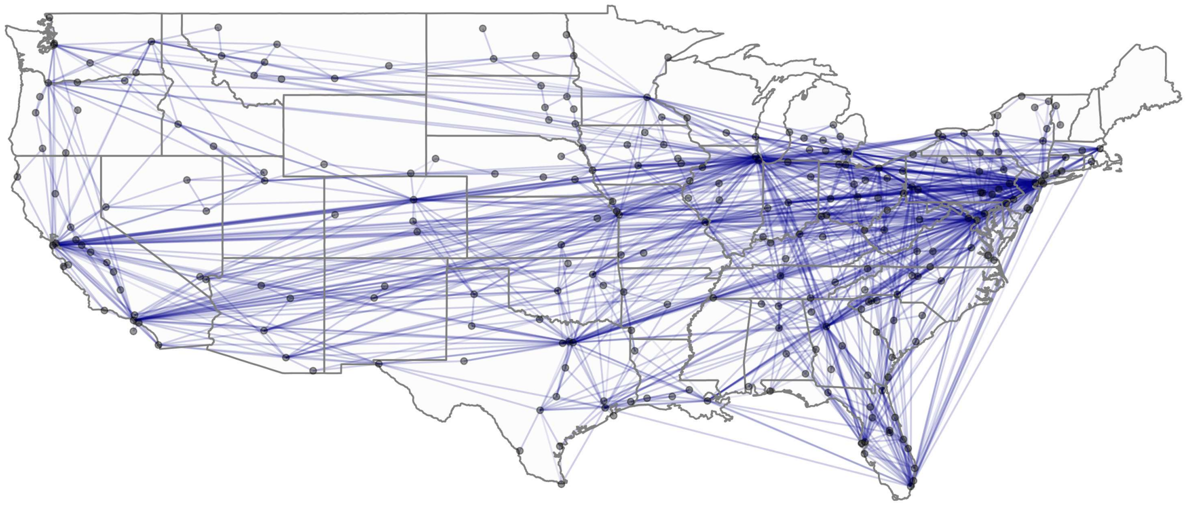


Fig. 1.4. United States flight network 1951-1966

inside the MSA or is at most 15km away from its boundary.[15] 176 out of 275 US airports are matched to at least one MSA. Meanwhile, 142 out of 168 MSAs are matched to one or more airports in at least one year, and 108 MSAs are matched to one or more airports in the four years. We use the sample of 108 MSAs that had an airport in the four years as our baseline travel time data.[16]

4.1. *Descriptive statistics: Air travel*

To understand the changes in travel time we will first study travel time of non-stop flights and then of all routes including connecting flights. Figure 1.5 displays the non-stop fastest flight within each MSA pair that was operating in each year. In 1951 the longest non-stop flight across MSAs was between Chicago and San Francisco using the Douglas DC-6, covering a distance of 2,960 km in 7 hours 40 minutes. This travel time was just under 8 hours, the maximum flight time allowed for a crew in a 24-hour period.[17] In 1956, new regulation allowed up to 10 hour flights for transcontinental flights, the longest non-stop

[15]The 15km distance was chosen after inspecting airports near the border that should arguably be matched, as for example, Atlanta ATL airport.

[16]In Appendix A.2 we include a table with the 168 MSAs, those connected at least once and those connected in the four years. Among the MSAs not connected is San Jose, California, which in our patent sample accounted for around 2% of patents. San Jose had an airport (SJC) during our time period but it was not served by any of our airlines, so it is not included in our analysis. In the future we plan to include the currently non-connected MSAs by matching them to airports that may have served them and accounting for the commuting travel time.

[17]Honolulu was not concerned by the regulation. Honolulu was connected with non-stop flights to San Francisco (9 hours 40 minutes), Los Angeles (11 hours) and Portland (12 hours 55 minutes).

flight between MSAs was New York to San Francisco with the Douglas DC-7, covering a distance of 4,151 km in 9 hours. Between 1951 and 1956, while we observe an increase in average flight speed that went up to 17%, the main change observed is that longer non-stop routes were possible.

In 1961, the first year in which we have jet aircrafts in the travel time data, there is a reduction in travel time between MSA-pairs, especially for those far apart from each other. In 1966, there is a further decrease in travel time due to a widespread adoption of jet aircrafts in shorter distances. In Appendix Figure 1.22 we show the jet adoption rate by distance for MSAs connected with a non-stop flight. All MSA-pairs more than 3,000km apart connected with a non-stop flight operated at least one jet flight in 1961, and this expanded to all those more than 2,000km apart in 1966. The speed gain of jets relative to propeller aircrafts is increasing with the amount of time that the jet can fly at its cruise speed, arguing in favor of an adoption that is increasing with the distance between origin and destination.[18]

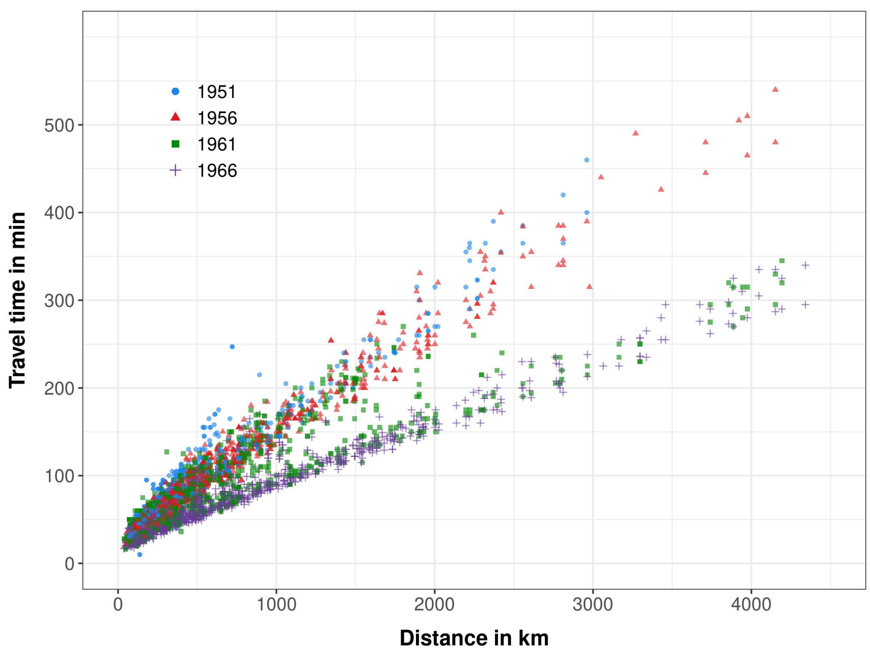


Fig. 1.5. Non-stop fastest flights United States MSAs

The change in travel time in non-stop flights is also reflected in the travel time for connect-

[18]We are currently exploring the differential timing of jet adoption across airlines. Differences in (pre-existing) route distance and past contractual relationships with aircraft suppliers potentially led to different adoption rates at each time period. For example, Eastern Airlines' routes were particularly shorter than for other airlines. Also, those committed to buy Douglas airplanes (the leader US commercial aircraft supplier pre-jet era) would have adopted jets later, as Douglas launched jet airplanes later than Boeing.

ing flights. Figure 1.6 shows, relative to 1951, the average and standard deviation change in travel time for all MSA-pairs, including non-stop and connecting flights.[19] Between 1951 and 1956, there is an average reduction in travel time of 9.2% which is roughly constant for all distances over 500km. Between 1951 and 1961, there is a reduction in travel time that is increasing with distance. The average decrease in travel time is of 16.8%, while the reduction is of 29.4% for a distance of more than 2,000km and 39.2% for a distance of 4,250-4,500km. Between 1951 and 1966, there is an even stronger decrease in travel time at all distances. The average reduction in travel time is 28.7% across all distances, 40.8% for a distance of more than 2,000km and 48.4% for a distance of 4,250-4,500km. The increased adoption of jets for short distance flights implied that both non-stop flights at short distance and connecting flights at farther distance had a decrease in travel time.

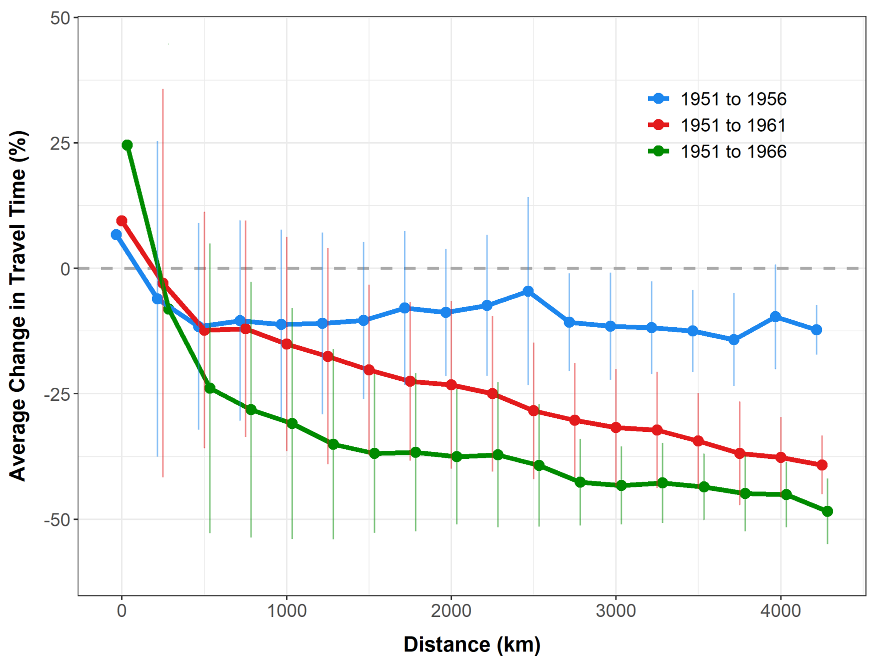


Fig. 1.6. Change in MSAs travel time

Figure 1.25 in Appendix A.2 shows that the change in travel time is accompanied by a reduction of the amount of legs needed to connect two MSAs at every distance. This reduction is especially marked between 1951 and 1956, and 1961 and 1966. Between 1956 and 1961, we do not observe a big reduction in the amount of legs, implying that the decrease in travel time observed in Figure 1.6 between 1956 and 1961 comes from a source other than the amount of legs. In Appendix Figure 1.26 we open up the change

[19]The plot includes only MSA-pairs with travel time in all time periods. The standard deviation for MSA-pairs less than 250km apart is big relative to the ones at other distances. Hence we decided not to include it because it distorts the visualization of the rest of the plot.

in travel time by the way an MSA pair was connected in 1951 and 1966: either directly (non-stop flight) or indirectly (connecting flight). We observe that much of the increase in travel time for MSA pairs less than 250km apart comes from routes that in 1951 were operated non-stop while in 1966 were operated with connecting flights.[20] Interestingly, for MSA-pairs more than 2,000km apart travel time reduced on average 42% for those pairs that were connected indirectly in both periods, and 51% for those that switched from indirect to direct. This fact shows the relevance of improvements in flight technology even for MSAs that were not directly connected.

It could be the case that a reduction in the amount of legs or an increase in frequency of flights reduces layover time, which then translates into a reduction of travel time. In Appendix Figure 1.28 we compare the change in travel time from 1951 to 1966 with a counterfactual change in travel time in which we eliminate layover time in both time periods. We observe that the average change in travel time is stronger at every distance in the counterfactual scenario without layover time. This implies that the relative importance of layover time to total travel time within a route increased between 1951 and 1966, so total travel time did not decrease proportionally to the change of in-flight travel time. In short, layover time attenuated the reduction in travel time.

4.2. Constructing an instrument

In this section we construct an instrumental travel time that is based on the pre-existing flight routes and the time-varying nationwide roll out of jets. In this way, the instrument abstracts from the endogenous decisions of two agents: First, regulator's decision on the opening/closure of routes. Second, airlines' decision about to which routes allocate jet vs propeller airplanes and scheduling (frequency of flights and layover time). We first explain the idea and identifying assumptions of the instrument, and then we detail how it is constructed.

In Borenstein and Rose (2014) it is argued that, due to strict regulation, it was difficult for airlines to adapt their flight network when technology to fly changed. However, we may be concerned that the decision of the regulator to grant new routes could be targeted to specific pairs or correlated with unobservable variables that also affect the creation and

[20] Appendix Figure 1.27 repeats the exercise discarding layover time in all time periods. By comparing Figure 1.26 and Figure 1.27 we can disentangle the effect of layover time and the change in in-flight time. For MSA pairs less than 250km that changed from direct to indirect connection, 80% of the increase in travel time is due to the increase in layover time (which was previously zero as it was a non-stop flight), and 20% is due to the increase of in-flight time.

diffusion of knowledge.[21] Hence, as the first step in the construction of our instrument, we *fix routes* to the ones we observe in 1951. In this way the instrumental travel time is computed only using non-stop flights present in 1951, and does not consider appearance or disappearance of non-stop flights in the data. The identifying assumption is that the network of flight routes in 1951 did not yet include the changes that would be optimal to operate with jet airplanes. In other words, we require that the regulator did not change routes already by 1951 in anticipation of the arrival of jet airplanes.[22,23]

Airlines could decide on two factors that affect travel time: the type of airplane (jet vs. propeller) operated in each route and scheduling, which consists on the frequency of flights and layover time in case of connecting flights.[24] We may be concerned that, as with the regulator, airlines' decisions could be correlated with unobservables that also affect the creation and diffusion of knowledge.[25] The second step in the construction of our instrument is to discard layover time (hence discarding all scheduling decisions) in all time periods, and assume that in each year all routes are operated with a *fictitious average airplane* of the year. Hence, the change in instrumental travel time in a route is independent of the type of airplane used in the route and it only depends on the nationwide roll out of jets. The identifying assumption is that no single route had the power to shift the average speed of the year.

To construct the instrumental travel time we first estimate, separately for each year, a linear regression of travel time on flight distance using only the fastest non-stop flight in each origin-destination airport pairs.[26] These yearly regressions provide us with the fictitious average airplane of each year: the intercept gives the take-off and landing time of the airplane while the slope provides the (inverse) speed. Second, we fit these regressions to obtain predicted travel time in each non-stop flight and year. Third, for each year, we

[21] For example, the regulator could have targeted the opening of new routes between places in order to boost their economic activity.

[22] For example, in the instrument there are no non-stop transcontinental routes.

[23] In our estimations we exploit time variation for identification. Hence, if pre-existing routes affect the levels at the origin-destination level, this does not drive our identification. However, we may be concerned that pre-existing routes could affect future growth and not only levels. To address this concern, in robustness analysis we estimate the elasticity of citations to travel time using only MSA-pairs that are always indirectly connected. Results go in the same direction.

[24] In 1961, all non-stop flights of more than 3,000km had at least one jet operating within them, while in 1966 it was the case in all non-stop flights of more than 2,000km. Therefore the endogeneity of jet adoption is a smaller concern for long distance flights.

[25] For example, airlines may have decided to prioritize the allocation of jets to routes which had a higher share of business travel, which may be correlated with the diffusion of knowledge.

[26] The use of a linear regression is motivated by the linearity between travel time and distance displayed in Figure 1.5. To estimate these regressions we use all routes appearing in each year.

compute the fastest travel time using the Dijkstra algorithm. The Dijkstra algorithm looks for the fastest path using only 1951 non-stop flights, while the travel time in each non-stop flight in each year is given by the predicted travel time from the previous step. Layover time is set to zero in all years.

Figure 1.7 shows the percentage change in observed and instrumental travel time relative to 1951. We compute the percentage change within each MSA-pair for each year and then take averages within 250km bins. We observe that the instrumental travel time follows pretty closely the observed change in travel time in each year. Especially, it replicates the pattern of a stronger decrease in travel time for MSAs located farther apart. It is only for MSAs less than 250-500km apart that the change in the instrumental travel time departs from the observed change.[27] This finding shows that most of the change in travel time that we observe is due to the change in speed of airplanes, and that the endogeneity concern is limited for MSAs located far away from each other.

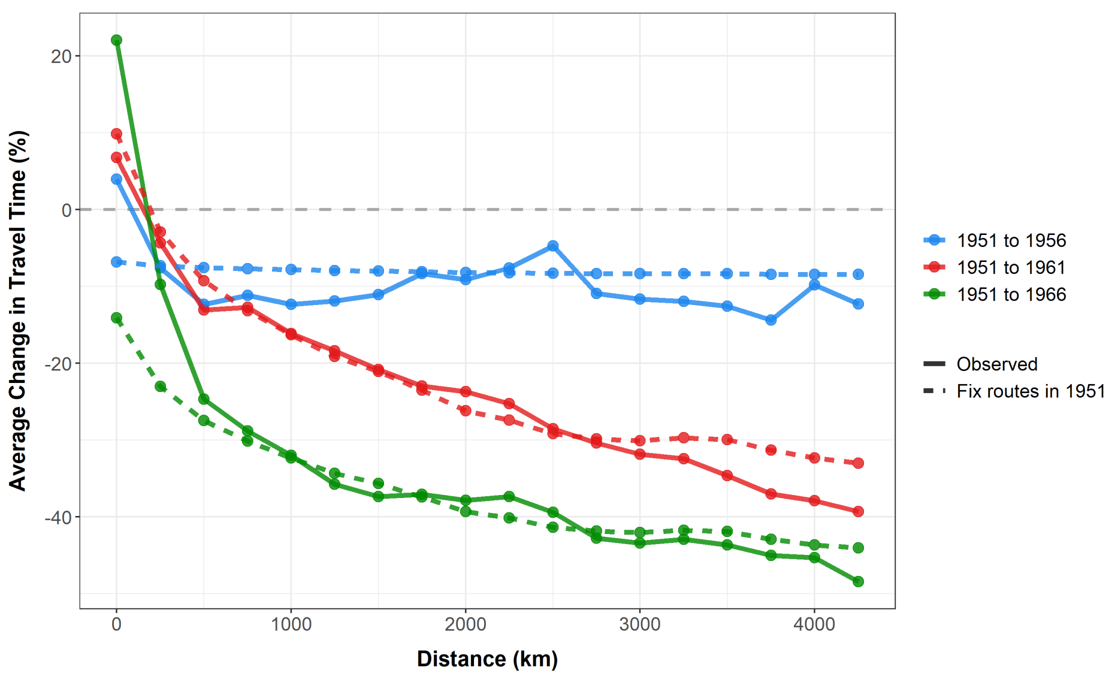


Fig. 1.7. Instrumental Travel Time between US MSAs.

In Appendix A.2 we present other two counterfactual travel times: one in which we

[27]We observe an increase in travel time for short distances in 1961 relative to 1951. Given that non-stop routes are fixed and that for longer distances there is a decrease in travel time, the increase in travel time in short distances comes from an increase in the value of the intercept relative to the slope in 1961, relative to 1951.

fix airplanes to be the average airplane of 1951 and allow routes to evolve, and another in which both the average airplane and routes are varying. These two counterfactuals together with the one presented in this section allow us to decompose the change in travel time by the change in routes and the change in speed of airplanes. We obtain that around 90% of the change in travel time is due to the change in speed of airplanes, while around 10% of the change is due to the change in the flight routes. Appendix Figure 1.30 shows that the share is roughly constant for all distances. This finding confirms that most of the observed changes in travel time are due to improvements in flight technology.

5. Patent data

We use patent data to measure innovation. We construct a dataset of all patents granted by the United States Patent and Trademark Office (USPTO) with filing year[28] between 1949 and 1968, which includes for each patent: filing year, technology classification, location of the inventors when they applied for the patent, owner of the patent and citations to other patents also granted in the United States. This dataset provides the distribution of patents and citations over the geographic space, allowing to take into account ownership structure.

To construct the patent dataset we downloaded from Google Patents all patents granted by the USPTO with filing year between 1949 and 1968. This dataset contains patent number, filing year and citations.[29,30] Based on the patent number we merge it with multiple datasets. First, we obtained technology class from the USPTO Master Classification File[31] and we aggregated them to the six technology categories of Hall et al. (2001). Second, we obtained geographic location of inventors from three datasets: HistPat (Petralia et al. (2016)) and HistPat International (Petralia (2019)) for patents published until 1975, Fung Institute (Balsmeier et al. (2018)) for patents published after 1975.[32] We match all inventors'

[28]Filing year, also called application year, is the closest date to the date of invention that is present in the data and it represents the date of the first administrative event in order to obtain a patent. In the other hand, the publishing (also called granting year) is a later year in which the patent is granted. The difference between filing and publishing year depends on diverse non-innovation related factors (as capacity of the patent office to revise applications) and changes over time. Hence filing year is the date in our data that approximates the best to the date of invention.

[29]Very few patents had missing information on filing year. We complemented both missing filing year and citations with the OCR USPTO dataset.

[30]We note that the patent citation record starts in 1947, year in which the USPTO made it compulsory to have front page citations of prior art. Gross (2019)

[31]`https://www.google.com/googlebooks/uspto-patents-class.html`

[32]Due to the gap between the filing year and publishing year we also do the matching to patents published

locations to 1950 Metropolitan Statistical Areas (MSAs) in contiguous United States. To do the match we obtain geographical coordinates from the GeoNames US Gazetteer file and Open Street Maps, and use the MSAs shape file from Manson et al. (2020). Third, we obtain ownership of patents from two sources: Kogan et al. (2017) for patents owned by firms listed in the US stock market and Patstat (Magerman et al. (2006)) for the remaining unmatched patents.[33]

For the descriptives presented below and the posterior analysis we truncate and aggregate the data in the following way. We drop patents that are owned by universities or government organizations. To count patents that are classified into multiple technology categories, we do a fractional count by assigning proportionally a part of the patent to each category. Citations are counted as the multiplication of the technology weight of the citing and cited patents. We drop patents (and their citations) that have inventors in multiple MSAs[34] and citations in which the citing owner is the same as the cited owner.[35]

We aggregate the patent data to 4 time periods of 5 years each, with the center of each period being the year of travel time data collected. The periods are: 1951 (which contains the years 1949-1953), 1956 (1954-1958), 1961 (1959-1963) and 1966 (1964-1968). We consider only patents in MSAs that are matched to an airport in the four periods.[36] The final dataset contains 108 MSAs with patents and travel time.

5.1. *Descriptive statistics: Patents*

This section presents three facts about US patents over our sample period: First, initially less innovative locations had a higher patenting growth rate. The average yearly growth rate of locations in the lowest quartile of initial innovativeness was 7.2% while it was 1.9%

after 1968. Our underlying patent data actually covers a longer time period of filing years, which we need for example to construct forward and backward citation lags. However, there are limitations to use the geographic data in filing years 1971-1972. In Appendix B we show that during filing years 1971-1972 the rate of unmatched patents to inventors' location increases. This is probably due to Histpat and Fung data not being a perfect continuation one of the other.

[33]Patent ownership in both datasets comes from the patent text, which is self declared by the patent applicant. Particularly, Kogan et al. (2017) does not explicitly state if it takes into account firm-ownership structure to determine the ultimate owner of a patent, neither does Patstat.

[34]Working with multi-MSA patents requires an assumption on how to compute distance and travel time between the citing and cited patents, as it is not a single origin-destination location pair. We hence prefer to abstract from multi-MSA patents. Nevertheless, collaboration of inventors located in different MSAs is an interesting subject for further research.

[35]Incentives to self-cite may be different than to cite patents of other owners.

[36]We drop around 9% of patents that are in MSAs which are not matched to an airport in the four time periods. Descriptive statistics including those patents are similar to the ones presented here.

for those in the highest quartile. High growth locations were also primarily in the South and the West of the US. The South and the West grew three times as fast as the Midwest and the Northeast. Second, over time firms grew larger as measured by the amount of MSAs in which they had research establishments. At the same time, the share of patents filed by large multi-establishment firms increased. The amount of firms with research establishments in more than 10 MSAs almost tripled over the time period and their share of patents doubled. Third, the mass of citations shifted towards longer distances. While the first quartile of citation distance remained relative stable over the time period, the third quartile increased its distance by 39%. At the same time, the share of citations at more than 2,000km increased by 30%.

Below we present descriptive stastistics averaging across technologies. Technology-specific descriptives are included in Appendix B.

Fact 1.a.: Initially less innovative locations had a higher patenting growth rate

In the period 1951 to 1966 we observe that the highest growth of patenting takes place in locations that were initially less innovative. The differential growth rate implies a convergence rate of 5.3% per year.

Figure 1.8 shows the geographic distribution of patenting in 1951. Darker colors refer to a higher level of *initial innovativeness*, which is defined as the amount of patents filed by inventors in the MSA in 1951.[37] We observe that MSAs in the top quartile of patenting are concentrated in the Northeast (which includes New York) and the Midwest (which includes Chicago), with few additional MSAs in the West.[38,39]

Figure 1.9 shows the geographic distribution of patenting growth in 1951-1966.[40] We observe a striking pattern relative to Figure 1.8: high growth MSAs were those that were

[37]To compute the level of initial innovativeness we only use patents filed in 1951 (years 1949-1953). We aggregate patents to the MSA-technology level and then compute the quantile-position of each MSA in the technology. Lower values of quantile-position refers to lower amount of patents in the technology (relative to other MSAs). Each MSA has a different value of quantile-position in each of the 6 technology categories. To obtain the MSA level quantile we take the average quantile across technologies within the MSA. Finally we classify MSAs into quartiles depending on whether the average quantile is higher or lower than the thresholds 0.25, 0.50, 0.75.

[38]In Appendix B we show that the 1951 geographic distribution of patents looks similar across technology categories.

[39]The top 5 patenting MSAs in 1951 were: New York City (25% of all patents), Chicago (11%), Los Angeles (8%), Philadelphia (6%) and Boston (4%).

[40]We compute the growth rate of patenting in each technology within a MSA and then take the average across technologies within the MSA.

initially less innovative. High growth happens in initially less innovative locations in the South and the West but also in the Northeast. We confirm this pattern in Figure 1.10, which shows the MSA's ranking of innovativeness in 1951 and its subsequent patenting growth rate in 1951-1966. Figure 1.10 shows that MSAs that were initially more innovative (lower values in the ranking) are those that saw lower values of subsequent patenting growth.[41,42] We estimate a linear regression with an intercept and a slope, and find that the slope is positive and statistically different from zero. At the mean, lowering initial innovativeness by 10 positions in the ranking was associated with a subsequent 0.42 percentage points higher yearly growth rate of patenting.

Figure 1.10 presents average growth rates across technologies within a MSA. We obtain a result that goes in the same direction if we compute the average growth rates across MSAs within a technology and quartile of initial innovativeness, and then take the average across technologies. The average yearly growth rate of MSA-technologies in the lowest quartile of initial innovativeness is 7.2% while it is 1.9% in the highest quartile.[43] The percentage point difference between the two growth rates implies that locations in the lowest quartile converged towards locations in the highest quartile at a speed of 5.3% per year.[44] The convergence in patenting across MSAs is consistent with *The Postwar Decline in Concentration, 1945-1990* described in Andrews and Whalley (2021).

[41] Each dot in Figure 1.10 is an MSA. To compute the MSA ranking we need to double-rank MSAs. First we rank all MSAs in each technology. Second we take the across-technology average ranking of each MSA. Third we rank all MSA's averages. To compute the MSA's yearly growth rate we first take the 1951-1966 growth rate for each technology in the MSA. We then take the average across technology. Finally we obtain the MSA's yearly growth rate by computing: $yearly\_growth\_rate = (1 + 19\_year\_growth\_rate)^{(1/19)} - 1$ (the 1951 to 1966 period is a 20 year window, we take growth rates as being from the first year 1949 to the last one 1968, which is 19 year growth).

[42] In Appendix B we show replicate the plot differentiating geographic regions. MSAs that were initially less innovative and had high subsequent growth were located in all four regions, although they were primarily located in the South and the West.

[43] We first compute the 1951-1966 growth rate (19-year growth rate) for each MSA-technology. We then take averages across MSAs within a quartile-technology, and after take averages across technologies within a quartile. Finally, we convert the 19-year growth rate into an average yearly growth rate.

[44] We note that the aggregate growth of patents is much smaller than the across MSAs unweighted average, and this is exactly because initially less innovative MSAs grew faster. If we compute the growth rate in nationwide amount of patents in each of the technologies and then average across technologies we obtain a yearly growth rate of 1.5%.

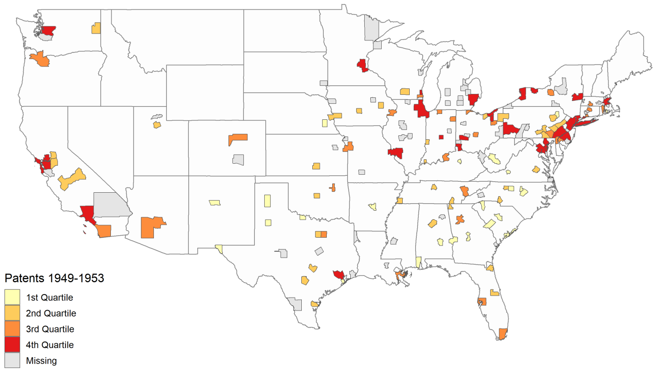


Fig. 1.8. Geography of Patenting 1951

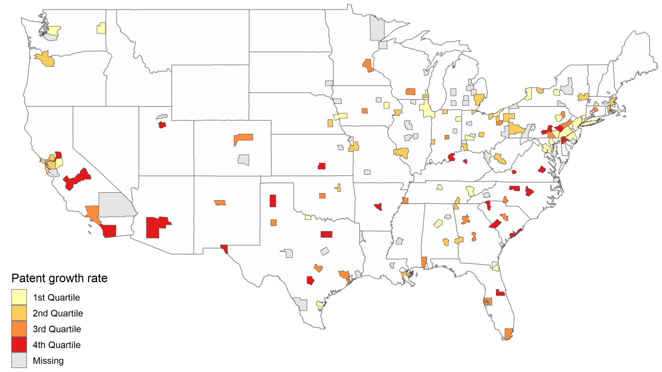


Fig. 1.9. Patent growth 1951-1966

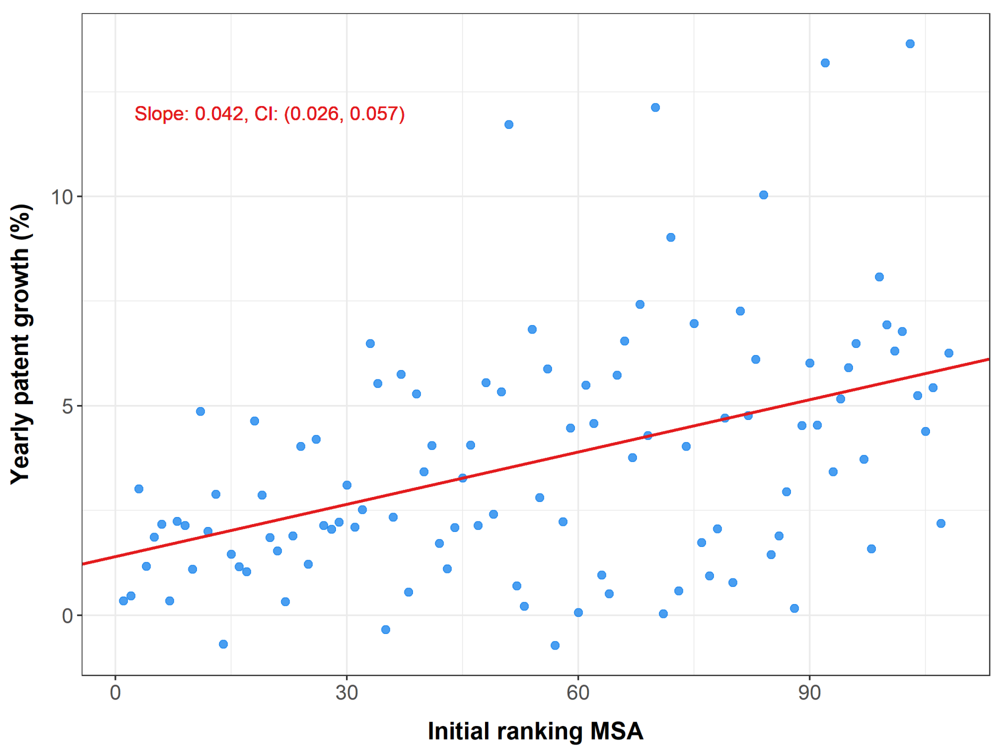


Fig. 1.10. Patent growth rate by initial innovativeness ranking of MSA

Fact 1.b.: The South and the West of the US had a higher patenting growth rate

Figure 1.9 shows that MSAs located in the South and the West of the US had a higher patenting growth rate in 1951-1966. We classify MSAs using Census Regions of the

US (Midwest, Northeast, South and West)[45] and aggregate patents within each region-technology-year. Figures 1.11 and 1.12 present averages across technologies within a region-year. Figure 1.11 shows that the share of patents filed by inventors located in the Midwest and the Northeast decreased from 75% in 1951 to 68% in 1966, while the share of patents filed in the South and the West increased from 25% to 32%. The opposite change in the shares implies that the South and the West had a higher growth rate of patenting relative to the Midwest and the Northeast.

Figure 1.12 shows that in the period 1951-1966 the South and the West increased their amount of patenting by 80%, while the Midwest and the Northeast had a 22% growth.[46] Translated into yearly growth rates, the South and the West grew three times as fast as the Midwest and the Northeast (3.14% vs. 1.05% per year).[47]

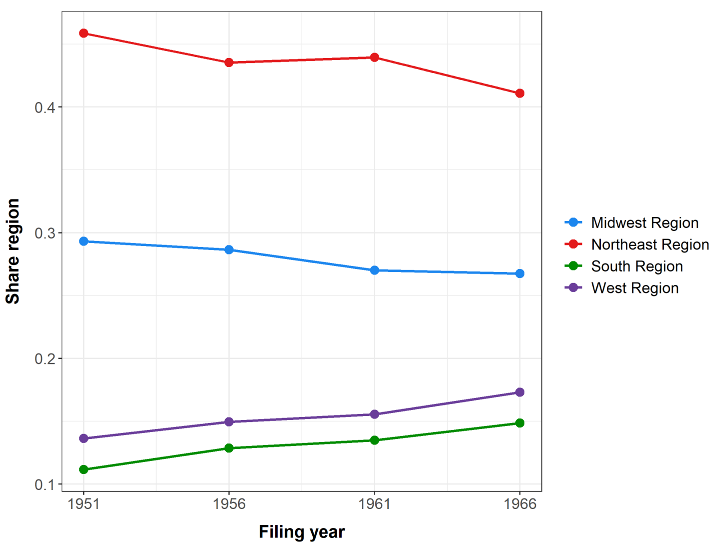


Fig. 1.11. Share of patents by region

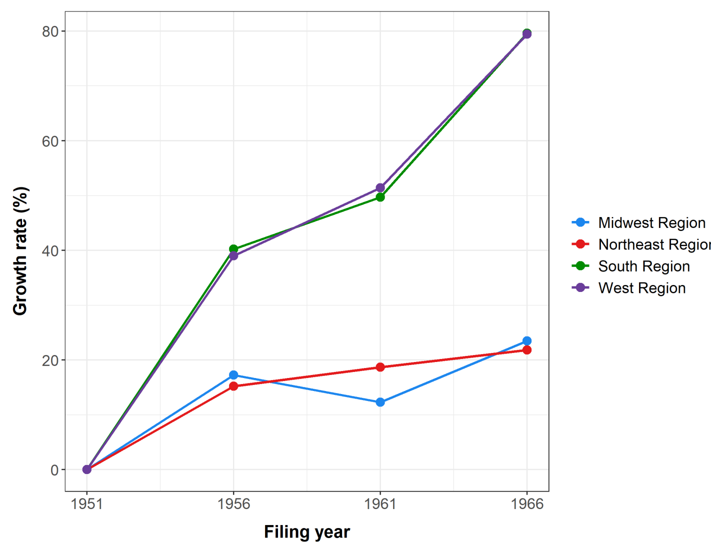


Fig. 1.12. Patent growth by region

Fact 2: Multi-establishment firms expanded geographically and accounted for a higher share of patents

Using all the patents of the same owner we identify all locations in which a patent owner had inventors applying for patents. We label a patent owner a *firm* and assume that a firm has a *research establishment* in the MSAs in which it has inventors applying for patents. Combining all patents belonging to the same firm we know if a firm has research

[45] In Appendix C we present a map with the four Census Regions.
[46] Growth rates are computed by region-technology and then averaged across technologies within region.
[47] $3.14\% = 1.80^{(1/19)} \times 100$, $1.05\% = 1.22^{(1/19)} \times 100$

establishments in multiple MSAs, if a firm expands over time and where it locates its establishments.

In Table 1.1 we count the number of firms and compute their share of patents according to whether the firm had 1, 2 to 5, 6 to 10, 11 to 20, or more than 20 establishments in each respective year. As we can see, the vast majority of firms had one establishment (95.8% in 1951), while very few had 11 or more establishments (0.1% in 1951). In 1951, single-establishment firms accounted for 57% of all patents. At the same time, firms with 11 or more establishments (42 firms, 0.1% of all firms) accounted for 15% of all patents.

From 1951 to 1966, the amount of single establishment firms declined by 1% while the amount of firms with 11 or more establishments increased by 283%. In other words, the amount of firms with presence in 11 MSAs or more grew from 42 to 119 firms. At the same time, the share of patents accounted by firms with 11 or more establishments increased from 15% to 31%. Simultaneously, the share of patents of single-establishment firms decreased from 57% to 46%. Hence, Table 1.1 illustrates that both the amount of multi-establishment firms and their share of patents grew over time.[48] In Appendix B we show that multi-establishment firms increased their share of patents in all quartiles of MSAs' initial innovativeness, with a stronger increase in initially less innovative MSAs.

Table 1.1: Number of firms and share of patents by firm's geographic coverage

	Number of firms					Share of patents				
Year \ N. estab.	1	2 to 5	6 to 10	11 to 20	+20	1	2 to 5	6 to 10	11 to 20	+20
1951	41,133	1,684	75	34	8	0.57	0.19	0.08	0.07	0.08
1956	42,590	1,927	111	60	12	0.52	0.19	0.09	0.11	0.08
1961	37,366	2,112	131	80	18	0.48	0.19	0.09	0.13	0.12
1966	40,711	2,086	132	89	30	0.46	0.15	0.09	0.14	0.17

Geographic coverage is computed as the amount of Metropolitan Statistical Areas (MSAs) in which the firm has inventors applying for patents (*research establishments*) in a certain year. Bins of geographic coverage are 1 MSA, 2 to 5 MSAs, 6 to 10 MSAs, 11 to 20 MSAs, more than 20 MSAs. The maximum possible is 108 MSAs.

While we observe an increase in the number of multi-establishment firms, we also observe an increase in the distance between establishments of the same firm. Figure 1.13 shows

[48] Within each year and bin of firm size, we compute the share of patents by technology and then take the average across technologies. We have computed the across-firms Herfindahl index within technology (so it shows the level of across-firm concentration within a technology) and we do not observe a clear pattern of either concentration or deconcentration.

that, for firms that have multiple establishments in the respective year, the average distance across establishments within the firm increased over time.[49]

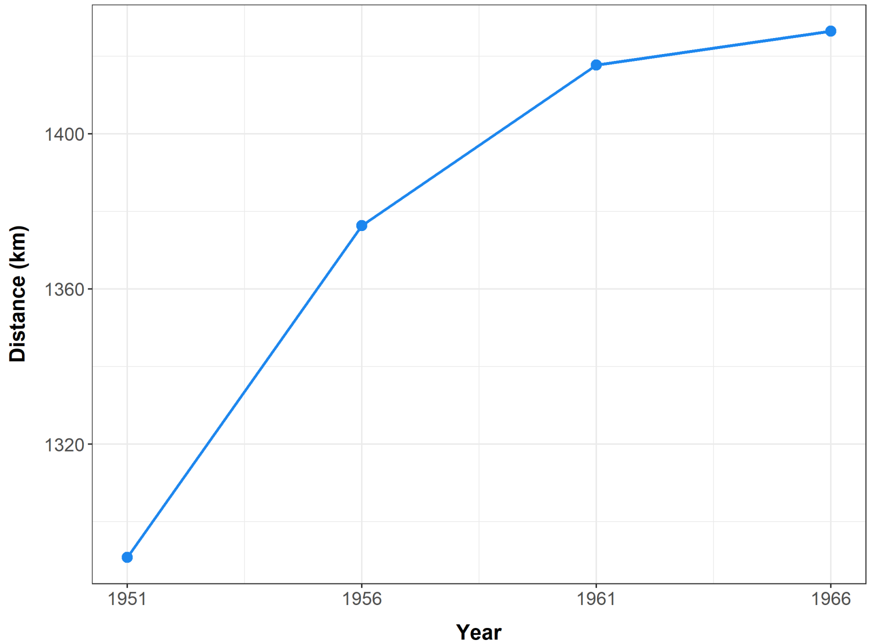


Fig. 1.13. Average distance across establishments within the firm

Fact 3: Distance of citations increased

In our analysis we use citations as a proxy for knowledge diffusion. According to Jaffe et al. (1993) *"a citation of Patent X by Patent Y means that X represents a piece of previously existing knowledge upon which Y builds."* (page 580).[50] We compute the distance between the citing inventor and the cited inventor. Figure 1.14 shows the evolution over time of the first, second and third quartile of citation distance.[51] We observe that 25% of citations happened between inventors located less than 300km apart throughout our sample period. For the middle 50% of citations we observe that over time inventors cited other inventors located farther away. The third quartile of citation distance increased from 1,642km in 1951 to 2,284km in 1961, a 39% increase in the distance.[52] In other words, the mass of citations

[49]The increase in distance across establishments within firms could well be the result of firms that are growing and randomly producing new patents in different locations. However, in Section 8 we show that the process firms' geographic expansion was not random: firm's expansion was directed towards locations that got larger reductions in travel time to the firm's headquarters.

[50]Jaffe et al. (1993) discusses the reasons why to cite and why not to cite. Using a survey of inventors, Jaffe et al. (2000) find that there is communication among inventors and citations are a *"noisy signal of the presence of spillovers."*

[51]We compute distance between MSA centroids.

[52]As a reference, the distance from New York City NY to other places is: Boston MA 300km, Chicago IL 1,140km, Dallas TX 2,200km, San Francisco CA 4,130km. The quantile 0.10 of was at 0km in every period,

shifted towards longer distances.

In Figure 1.15 we present the share of citations by distance range between the citing and cited inventors.[53] The distance cutoffs where chosen in order to have a balanced share of citations in the initial time period, and considering the changes in travel time presented in Section 4.1. The share of citations that happen between inventors located more than 2,000km apart grew from 21.5% in 1951 to 27.9% in 1966. The 6.4 percentage points increase represents an increase of 30% of the share of citations at more than 2,000km.

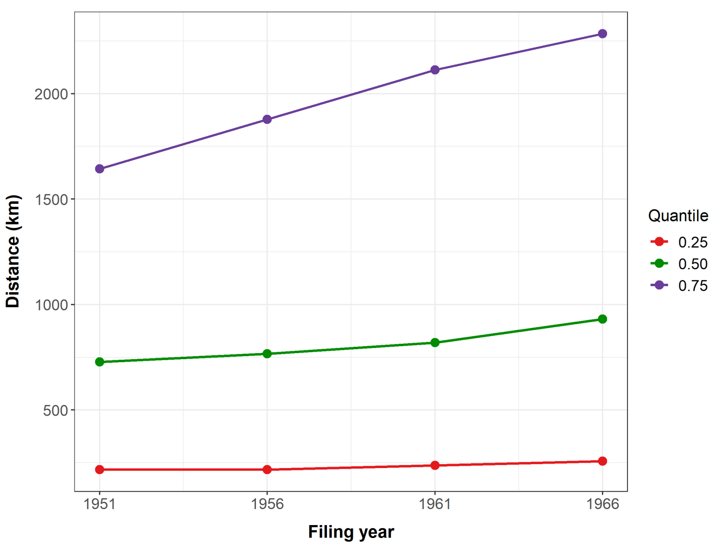


Fig. 1.14. Quantiles of citation distance

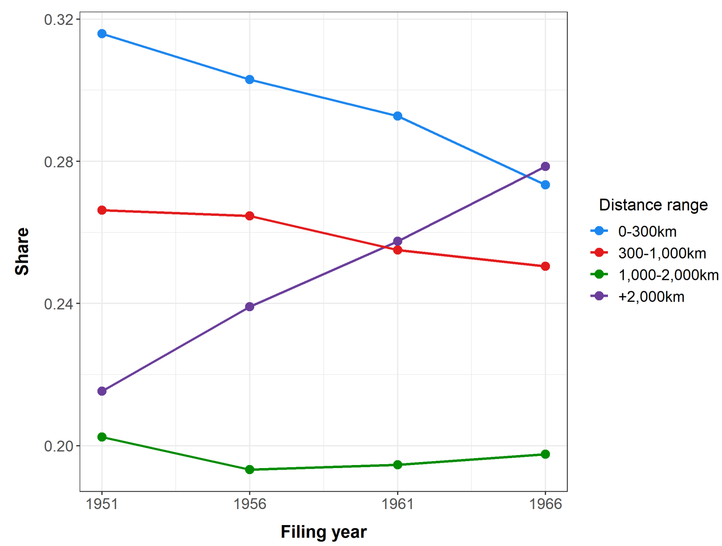


Fig. 1.15. Share of citations by distance

6. Diffusion of knowledge

In this section we show that the reduction in travel time led to an increase in knowledge diffusion, especially over long distances. In doing so we estimate the parameter β highlighted in equation (1.2): the elasticity of knowledge diffusion to travel time.

To perform the analysis we merge the air travel and patent datasets to obtain a final dataset that contains for each patent owner-location pair, the amount of patents filed in a

implying that 10% of citations took place within MSA. The quantile 0.90 was around 3,611km to 3,789km over the sample period.

[53]While Figure 1.14 shows how the distance of each quantile changes over time, Figure 1.15 shows the mass of citations (and hence the quantile to which belongs) in a certain distance cutoff. For example, in 1951 the share of citations in the 0-300km range was 31.6%, which is equal to saying that the quantile 0.316 in 1951 was 300km.

certain 5-year period and technology class, the amount of citations to other patents with their respective owner identifier, location and technology class, and the travel time to every location. We aggregate citations to the citing-cited establishment-technology within each period. We assume that passengers take a return flight, hence we make travel times symmetric.[54]

We estimate a gravity equation which relates citations between two establishments-technologies with their pairwise travel time.[55] We estimate the following regression:

$$citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt} \quad (1.3)$$

where $citations_{FiGjhkt}$ is the amount of citations from patents filed by the establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. We call Fi the research establishment of firm F in location i. travel time$_{ijt}$ is the air travel time (in minutes) between location i and j at time period t. The parameter of interest in the regression is $\boldsymbol{\beta}$, which represents the elasticity of citations to travel time.[56] If citations are affected negatively by travel time we would expect a negative value of β.

Given the panel structure of our data, we can include the fixed effect FE_{FiGjhk} that absorbs any time invariant citation behavior within the *citing establishment-technology and cited establishment-technology*. This fixed effect flexibly controls for persistent relationships within an establishment pair that would lead to relatively more (or less) citations. That includes characteristics like physical distance, but also pre-existing commercial relationships between establishments. The fixed effects FE_{Fiht} and FE_{Gjkt} control for the time changing general level of citations specific to each establishment and technology. For example FE_{Fiht} controls for the fact that if Fih files more patents in a given period, it would mechanically make more citations to every establishment. On the other hand, FE_{Gjkt} controls for Gjk filing more patents or higher quality patents that would receive more citations from every establishment.[57]

[54] $travel\ time_{ijt} = (travel\ time^{original}_{ijt} + travel\ time^{original}_{jit})/2$ where $travel\ time^{original}_{ijt}$ stands for the travel time between MSA i and j at time period t.

[55] For explanation and micro foundations of the gravity equation see Head and Mayer (2014) and references thereof.

[56] A 1 percent increase in travel time has an effect of β percent increase (or decrease in the case of a negative β) in citations.

[57] In the trade literature, the parallel of the fixed effects (simplified for exposition) would be: FE_{ij} country-pair fixed effect, FE_{jt} origin-time fixed effect and FE_{it} destination-time fixed effect.

The inclusion of FE_{FiGjhk} implies that only variation across time within an establishment-pair is used for identification. By additionally including the fixed effect FE_{Fiht}, the across-time variation is compared only between citing-cited establishment-technology pairs $FiGjhk$ within a citing establishment-technology Fih in period t. As we also include FE_{Gjkt}, the comparison is done while controlling for the size of the cited establishment-technology Gjk in period t. Put differently and simplifying slightly, the identification of $\boldsymbol{\beta}$ relies on changes in citations and travel time within an establishment-pair, relative to another establishment-pair with the same citing establishment, conditional on the two cited establishments' sizes.

Following Silva and Tenreyro (2006), we estimate the gravity equation by Poisson Pseudo Maximum Likelihood (PPML).[58] This estimation methodology has two advantages over a multiplicative model that is then log-linearized to obtain a log-log specification. First, it only requires the conditional mean of the dependent variable to be correctly specified, while the OLS estimation of the log-linearized model would lead to biased estimates in the presence of heteroskedascity. Second, it allows to include zeros in the dependent variable, which is especially relevant when using disaggregated data. One downside of estimating PPML with the fixed effects that we include is that both coefficients and standard errors have to be corrected due to the incidental parameter problem (Weidner and Zylkin (2021)). We follow Weidner and Zylkin (2021) to use split-panel jackknife bias-correction on the coefficients and Dhaene and Jochmans (2015) to bootstrap standard errors which we also bias-correct with split-panel jackknife.[59]

Whenever $FiGjhk$ has positive citations in at least one period and missing value in another, we impute zero citations in the missing period.[60] Travel time is set to one minute whenever $i = j$.[61]

[58]We use the package *fixest* (Bergé (2018)) in R to estimate high dimensional fixed effects generalized linear models *feglm* with Poisson link function.

[59]Details on the bias correction and bootstrap procedures are provided in Appendix D.

[60]We do not impute zeros in $FiGjhk$ that are always zero, as those observations would be dropped due to not being able to identify FE_{FiGjhk}.

[61]We measure air travel time in minutes. In our sample 13% of citations happen within the same MSA. The inclusion of those citations in the estimation increases the amount of observations available to identify of FE_{Fiht} and FE_{Gjkt}, and hence keeping them increases the amount of $FiGjhkt$ that remain in the effective sample to identify β. In order to include them we then need to impute a within-location travel time. We assume that within-location (air) travel time is not changing across time periods. Nonetheless, the identification of β is not affected by the value chosen for the within-location (time invariant) travel time, as β is identified by across time variation. In the appendix we show results using other values of (time invariant) within MSA travel time and the coefficients remain equal.

	PPML		IV PPML	
Dep. variable: *citations*		$citatitions_{FiGjhkt}$		
	(1)	(2)	(3)	(4)
log(travel time)	-0.083^{***} (0.019)		-0.152^{***} (0.029)	
log(travel time) × 0-300km		0.019 (0.036)		-0.076 (0.221)
log(travel time) × 300-1,000km		-0.089^{***} (0.023)		-0.134^{***} (0.044)
log(travel time) × 1,000-2,000km		-0.094^{***} (0.033)		-0.112^{**} (0.047)
log(travel time) × +2,000km		-0.169^{***} (0.039)		-0.203^{***} (0.043)
N obs. effective	$4,703,010$	$4,703,010$	$4,703,010$	$4,703,010$
R2	0.88	0.88	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.2: Elasticity of citations to travel time

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (2) includes the interaction of travel time$_{ijt}$ with a dummy for distance bin between the citing establishment Fi and the cited establishment Gj. Column (3) and (4) show the result of two step instrumental variables estimation, where $log(\text{travel time}_{ijt})$ is instrumented with $log(\text{travel time}_{ijt}^{\text{fix routes}})$, the travel time that would have taken place if routes were fixed to the ones observed in 1951 and in each year routes were operated with the average airplane of the year. Bootstrap standard errors are presented in parentheses. The coefficients and standard errors in columns (1) and (2) are jackknife bias-corrected. R2 is computed as the squared correlation between observed and fitted values.

Column (1) in Table 1.2 presents the results of estimating equation (1.3). The value of the elasticity of citations to travel time is estimated to be -0.083, statistically significant at the 1% level. Given the average reduction in travel time of 31.4% in the full estimating sample, the elasticity implies that citations increased on average 2.6% as consequence of the reduction in travel time. If we consider the average decrease in travel time across all MSAs in the baseline travel time data, the implied increase is 2.4%.[62]

The importance of air transport relative to other means of transport potentially depends on the distance to travel. Also, we observed in section 4.1 that the improvements in air travel time depended on the distance to travel, with a difference in jet adoption for travel distances under and over 2,000km. Taking these two characteristics into account, we estimate a variation of equation (1.3) in which we allow the elasticity of citations to travel time to vary by distance interval between the locations of citing and cited establishments.[63] Column (2) in Table 1.2 shows the result of this estimation.[64] The estimated value of the elasticity in absolute terms increases with distance, reaching -0.169 for distances of more than 2,000km. Between 1951 and 1966 the average change in travel time in the full estimating sample is 47.7% for a distance of more than 2,000km. The estimated elasticity implies that citations between establishments at more than 2,000km apart increased by 8.1% due to the decrease in travel time. In total citations at more than 2,000km increased by 21%, implying that the change in travel time can account for 38.2% of the observed increase. If instead we consider the 40.8% average reduction in travel time across MSAs in the raw data, the elasticity implies an increase in citations of 6.9%, accounting for 32.7% of the total citation increase.

In Appendix B we investigate different heterogeneous effects. We study how travel time affects the extensive margin of citations (whether an establishment cites another establishment or not) and the intensive margin (conditional on citing, how much it cites). We find the effect comes from both margins. We estimate an heterogeneous elasticity depending on the level of spatial concentration of the citing technology and the cited technology, we do not find a statistical difference. We also look at whether it is older patents or younger patents that get diffused, finding some slight evidence that it is technologies that take longer time to diffuse that increase more their diffusion with the reduction in

[62] These values come from the multiplication of the elasticity of citations to travel time 0.083 and the average decrease in travel time between 1951 and 1966: 31.4% in the full estimating sample and 28.7% in the raw data of travel time across MSAs.

[63] We compute distance between the geographical center of each MSA.

[64] The share of observations (citations) in each distance interval is: 0-300km 26.1% (28.5%), 300-1,000km 30.7% (28.5%), 1,000-2,000km 19.7% (23.4%), +2,000km 23.4% (19.6%).

travel time. We study citations to and from government patents, and self citations, on the whole we do not find a different pattern from the baseline. We also do not find a particular pattern of the elasticity depending on the citing *firm's size* as measured by the amount of patents filed in 1949-1953. Finally, we estimate the elasticity by citing and cited technology and most of the effect seems to come from when the citing and cited technologies are the same.

There are two types of threats to identification in estimating equation (1.3): (i) the potentially targeted changes in travel time, which could be due to the opening of new routes, the allocation of jets across routes, or changes in scheduling, and (ii) time changes in other variables at the MSA-pair level which also drive the diffusion of knowledge and are correlated with the changes in travel time. In the remaining of this section we address the first type of threat by estimating the model by instrumental variables. In the following subsection we address the second type of threat by adding multiple controls. In both cases we show that results do not change.

As mentioned in Section 4.2, we may be concerned that the timing and allocation of jets to routes and that the opening/closure of routes were not random. In case there is an omitted variable that drives both the change in travel time at the MSA pair level and the change in citations across establishments within the same MSA, we would estimate biased coefficients. In order to tackle the endogeneity concern due to omitted variables we do an instrumental variables estimation using the instrument proposed in Section 4.2. To implement the instrumental variables estimation we follow a control function approach described in Wooldridge (2014). We proceed in two steps estimating the following two equations:

$$\begin{aligned} log(\text{travel time})_{FiGjhkt} &= \lambda_2 \, log(\text{travel time}^{\text{fix routes}}_{FiGjhkt}) \\ &\quad + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} \, + \, u_{FiGjhkt} \end{aligned} \tag{1.4}$$

$$\begin{aligned} citations_{FiGjhkt} &= \exp\left[\beta \, log(\text{travel time}_{ijt}) \, + \, \lambda \, \hat{u}_{FiGjhkt} \right. \\ &\quad \left. + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \, \times \, v_{FiGjhkt} \end{aligned} \tag{1.5}$$

In a first step we estimate equation (1.4) and obtain estimated residuals $\hat{u}_{FiGjhkt}$. In a second step we use the estimated residuals as a regressor in equation (1.5) which *controls* for the endogenous component of travel time. To perform inference we bootstrap standard

errors.[65] According to Wooldridge (2014), there would be evidence of endogeneity if the parameter λ in equation (1.5) is estimated to be statistically different from zero.

Columns (3) and (4) of Table 1.2 show the results of the instrumental variables estimation. If airlines were allocating jet airplanes to routes that would have witnessed a higher degree of exchange of knowledge even in the absence of jets, then we would expect the instrumental variables estimate to be smaller in absolute terms relative to the baseline coefficient. On the other hand, if the regulator targeted the opening of new routes between places that were in a lower trend of exchange of knowledge, we would expect the instrumented coefficient to be larger in absolute terms. Column (3) estimates the elasticity to be -0.152, bigger in absolute value compared to the non-instrumented estimate. The instrumental variables corrects for a downward bias in absolute terms, which represents evidence in favor of the regulator targeting the opening of new routes between places that had a lower degree of exchange of knowledge.[66,67]

In column (4) of Table 1.2 we see the coefficients of the instrumental variable estimation by distance between the citing and cited establishments. We observe the presence of a bias in the same direction as in column (3), however the magnitude of the bias is smaller except for the distance bin 0-300km, which is not precisely estimated. In particular, at more than 2,000km, the coefficient is relatively similar to the baseline estimation. In Appendix E we show the regression including coefficients on the residual *controls*. If the coefficients on controls are statistically significant, that is evidence of endogeneity. While the control is statistically significant when using only one coefficient for all distances, none of them is statistically significant when opening the coefficient by distance range. In other words, we do not find evidence of endogeneity at long distances, especially at +2,000km.

The instrument used in the instrumental variables estimation is constructed using the 1951 flight network. We may be concerned that the 1951 flight network is correlated with

[65] Appendix D includes details on the bootstrap procedure.

[66] The incidental parameter problem is potential present also in the instrumental variables estimation (IV PPML). However, there is currently no bias-correction procedure available for IV-PPML that we are aware off. Hence, columns (3) and (4) in Table 1.2 are not bias-corrected. In column (2) of Table 1.3 we present the PPML estimation not bias-corrected.

[67] The literature on weak instruments for non-linear instrumental variables is scarce. The rule of thumb of Staiger and Stock (1997) based on the F statistic is constructed using the bias that a *weak instrument* generates in a linear second stage (see Staiger and Stock (1997), Stock and Yogo (2005) and Sanderson and Windmeijer (2016) for testing for weak instruments in linear IV regression). For informative purposes, in the first stage of the model estimated in column (3) in Table 1.2 we obtain $\hat{\lambda}_2 = 0.95$ with a standard error 0.039 (clustered at the non-directional location pair level, ij is the same location pair as ji), and a within R2 of 0.38 (the share of residual variation explained by the instrument, after projecting out fixed effects).

future changes of citations.[68] In order to address this concern in Appendix E we estimate equation (1.3) by restricting the sample to establishments in MSA-pairs that are always indirectly connected. Results go in the same direction.

6.1. *Diffusion of knowledge: Robustness*

We may be concerned that there are other variables that could drive the diffusion of knowledge and at the same time be correlated with the change in travel time. In order to bias the coefficients, such omitted variables should be time-changing at the origin-destination MSA pair and be systematically correlated with the change in MSA-pair air travel time.[69] We consider three potential variables that could bias our estimates: improvements in highways, improvements in telephone communication and changes in flight ticket prices. In Table 1.3 we show the results controlling for this variables separately, while in Appendix E we include them simultaneously. Estimates are robust to including these controls.

Columns (1) and (2) in Table 1.3 present the elasticity of citations to travel time by distance bin. In column (1) the elasticity is bias-corrected while in column (2) it is not.[70] We observe that not doing the bias correction does not qualitatively affect the results. Columns (3) to (6) include the additional controls and should be compared to column (2).

First, in 1947 the Congress published the official plan for the Interstate Highway System, a nationwide infrastructure plan to improve existing highways and build new ones (see Baum-Snow (2007), Michaels (2008), Jaworski and Kitchens (2019) and Herzog (2021)). In case the change in travel time by air is correlated with the change in travel time by highway, we would have an omitted variable bias if we include only one of them in the estimation. Taylor Jaworski and Carl Kitchens have graciously shared with us data on county-to-county highway travel time and travel costs for 1950, 1960 and 1970, which we converted to MSA-to-MSA and linearly interpolated to convert to the same years of our air travel data. Hence we have a MSA-to-MSA time-varying measure of travel time. In Appendix E we show the correlation of MSA-to-MSA change in air travel time and

[68] We include a establishment pair fixed effect in the regressions, so a potential correlation between the 1951 flight network and the level of citations between research establishments does not affect our estimation.

[69] Variables that are not time changing or that are time changing at the MSA or establishment level do not represent a threat to identification, as they are flexibly controlled for with the fixed effects.

[70] The jackknife bias-correction due to the incidental parameter problem is computationally intensive. Due to the computational burden, we have still not bias-corrected all estimations. Columns (2) to (6) of Table 1.3 do not include bias-correction.

highway travel time.

Second, other means of communication like telephone lines may have expanded or changed their price during the period of analysis. Haines et al. (2010) contains information on the share of households within each city with telephone lines in 1960. We aggregate the variable to the MSA level. For each MSA-pair, we take the log of the mean share of households with telephone lines.[71] To include the variable as control we interact it with a time dummy to make the measure time variant. The assumption behind the interaction is that, if telephone lines expanded or changed their price over the time period, this time-change specific to each year was proportional to the 1960 log mean share of the MSA-pair.

Third, during the period of analysis ticket prices were set by the Civil Aeronautics Board, so airlines could not set prices of their own tickets. Some airlines included a sample of prices in the last page of their booklet of flight schedules, which we digitized. In Appendix E we document multiple facts about prices. The relevant fact for this section is that during 1962-1963 we observe a drop in prices of around 20% for routes of more than 1,000km distance. We may be concerned that the change in flow of knowledge is actually consequence of the change in prices, which happens to be correlated with the change in travel time. Given that we do not have ticket prices for each route and year, we use an estimated route price which is time varying. We obtain estimated prices by using the sample of prices that we digitized and fitting, for each year, price on a third degree polynomial of distance between origin and destination. We use log of estimated prices as control.[72]

Column (3) to (5) of Table 1.3 include the described controls. Assuming the covariance across coefficients is zero, none of the coefficients is statistically different from the baseline coefficients either in column (1) or (2).[73]

Fourth, we control for a time-varying effect of distance on citations. We may believe that

[71] Data from the 1962 City Data Book which comes from the US Bureau of the Census. log(mean telephone share$_{ij}$ = log((telephone share$_i$+telephone share$_j$)/2). We take the log of the mean share because the share is a linear combination of origin MSA and destination MSA characteristics, hence perfectly explained by origin and destination fixed effects. Taking the log prevents this.

[72] In order to perform inference we should adjust standard errors by the fact that we have a predicted regressor as control variable.

[73] To perform a test of statistical difference across coefficients of different regressions we need to estimate the covariance between them. We are currently doing a joint-bootstrap to obtain the covariance and perform the test.

other variables may have an effect on the diffusion of knowledge, and those variables are related to the distance between the citing and cited establishments. In column (6) we include as control $log(distance)$ interacted with a time dummy. We observe that the coefficients reduce in magnitude, potentially due to the fact that the change in travel time is also correlated with distance, hence controlling for a time-varying effect of distance absorbs part of the effect. In spite of that, the coefficient for distance of more than 2,000km remains statistically significant at the 5%. This result shows that travel time and distance are not equivalent measures. Hence, it highlights the importance of the origin-destination time-varying travel time data when studying the impact of face to face interactions.

Finally, as we will see in section 8.2, the entry and exit of research establishments was not uniform across locations during the sample period. We may then be concerned that the change in diffusion of knowledge is only consequence of the change in the geographic location of innovation. In Appendix E we re-estimate equation (1.3) with different samples: first, using only citing establishments that were present in 1949-1953, and second using only citing and cited establishments that were present in 1949-1953. We find the coefficient at more than 2,000km remains comparable to the one in the baseline regression, statistically significant at the 1%.

7. Creation of knowledge

In this section we show that the reduction in travel time to innovative locations led to an increase in knowledge creation. We show that the effect on knowledge creation was stronger in initially less innovative locations, leading to convergence across locations in terms of innovation. Additionally, the reduction in travel time contributed to a change in the geographic distribution of knowledge creation, increasing the relative importance of locations in the South and the West of the United States.

We construct a measure of *Knowledge Access* by adapting equation (1.2) to an empirical set up with multiple technology categories and time periods. The measure of *Knowledge Access* (KA_{iht}) shows how *easy* it is in time period t for research establishments in location i and technology h to access knowledge created in other locations. We compute *Knowledge Access* as follows:

$$KA_{iht} = \sum_{k} \omega_{hk} \sum_{j, j \neq i} \text{Patent stock}_{jk,t=1953} \times \text{travel time}_{ijt}^{\beta} \tag{1.6}$$

	PPML bias-corrected	PPML not bias-corrected				
Dep. variable: *citations*	$citations_{FiGjhkt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time) × 0-300km	0.019	0.021	0.023	0.0198	0.025	0.032
	(0.036)	(0.039)	(0.039)	(0.039)	(0.038)	(0.040)
log(travel time) × 300-1,000km	-0.089^{***}	-0.099^{***}	-0.096^{***}	-0.094^{***}	-0.102^{***}	-0.075^{**}
	(0.023)	(0.027)	(0.028)	(0.027)	(0.027)	(0.030)
log(travel time) × 1,000-2,000km	-0.094^{***}	-0.093^{**}	-0.089^{**}	-0.071^{*}	-0.104^{**}	-0.040
	(0.033)	(0.042)	(0.044)	(0.042)	(0.042)	(0.052)
log(travel time) × +2,000km	-0.169^{***}	-0.185^{***}	-0.180^{***}	-0.172^{***}	-0.196^{***}	-0.124^{**}
	(0.039)	(0.049)	(0.050)	(0.050)	(0.049)	(0.059)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88	0.88	0.88
Controls:						
log(highway time)	-	-	Yes	-	-	-
log(telephone share) × time	-	-	-	Yes	-	-
log(price)	-	-	-	-	Yes	-
log(distance) × time	-	-	-	-	-	Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.3: Robustness: Elasticity of citations to travel time, Part 1

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \, \mathbb{1}\{distance_{ij} \in d\} \, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \, \mathbb{1}\{distance_{ij} \in d\} \, \mathbb{1}\{X_{FiGjhkt}\} \, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km]$, $(300km - 1000km]$, $(1000km - 2000km]$, $(2000km - max]$. Column (1) presents jackknife bias-corrected coefficients and bias-corrected bootstrap standard errors. Column (2) repeats column (1) without bias-correction. Relative to (2), columns (3) through (6) contain additional controls. Column (3) controls for log highway time between i and j at period t. Column (4) controls for the log of the mean share of households with telephone line in 1960 in ij pair interacted with a time dummy. Column (5) controls for log flight ticket price between i and j at period t. Column (6) controls for log distance ij interacted with a time dummy. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Columns (2) through (6) present standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

where, from right to left, travel $\text{time}_{ijt}^{\beta}$ is the travel time between locations i and j at time period t, to the power of the elasticity of diffusion of knowledge to travel time. Patent $\text{stock}_{jk,t=1953}$ is the discounted sum of patents produced in location j and technology k between 1941 and 1953.[74] ω_{hk} is the share of citations of technology h that go to technology k at the aggregate level in 1949-1953, similar to an input-output weight.[75] Then, KA_{iht} is a weighted sum of the patent stock in each other location and technology, where the weights are how easy it is to access that patent stock (travel $\text{time}_{ijt}^{\beta}$) multiplied by how relevant that knowledge is (ω_{hk}).

In order to reduce concerns of potential endogeneity of accessing knowledge and creating knowledge, we exclude the patent stock in the location itself from the sum (we only use $j \neq i$).[76]

The measure of *Knowledge Access* contains across-time variation within a location-technology ih, and cross-sectional variation across technologies h within a location i. The across-time variation is only due to the change in travel time between locations, every other component of the measure is fixed to its 1949-1953 level. The cross-sectional variation comes from a distribution of Patent $\text{stock}_{jk,t=1953}$ within k that is not uniform across j, and from the input-output weights ω_{hk}. The joint across-time and cross-sectional variation means that if travel time for ij reduces, there will be a differential change in *Knowledge Access* across h within i which depends on the initial patent stock and input-output weights.

The degree with which changes in travel time are reflected in access to knowledge depend on how *important* travel time is to get knowledge to diffuse, which is the elasticity of knowledge diffusion to travel time that we estimated in Section 6. As the baseline we use $\beta = 0.185$, which is the elasticity of citations to travel time at more than 2,000 km not bias corrected. In robustness we use distance-specific β and in Appendix E we do sensitivity analysis of the results to changing the value of β.

[74] Patent $\text{stock}_{jk,t=1953} = \sum_{y\in[1941,1953]} \text{Patents}_{jky} \times (1 - \text{depreciation rate})^{1953-y}$. We use a depreciation rate of 5%, which is in the range of average depreciation rates of R&D found by De Rassenfosse and Jaffe (2017). We decided to fix the patent stock and not to allow it to change over time, as changes in travel time will potentially lead to changes in patent stock creating a dynamic reinforcing effect between knowledge access and new knowledge. In this sense, we abstract from *dynamic* externalities that could be at play.

[75] $\omega_{hk} = citations_{hk,t=[1949,1953]} / citations_{h,t=[1949,1953]}$ is included to weight each *source* technology category k by how important it is for the *destination* technology category h.

[76] The theory makes no distinction on whether the knowledge stock is in i or j, so in principle we would like to include the patent stock of i in the knowledge access of i. However, this could lead to econometric problems. First, we do not have exogenous variation of travel time within i. Second, if knowledge creation in i is a persistent process, by including the patent stock of i we would introduce a mechanical relationship between knowledge access and knowledge creation. Hence, our baseline measure of knowledge access of i does not consider the patent stock of i. In Appendix E we show that the inclusion of i's patent stock does not affect the results.

The measure of *Knowledge Access* allows us to translate changes in travel time between pairs of MSAs into a single location-technology specific characteristic, and to represent it on the same scale as patent growth in Figure 1.9. Figure 1.16 depicts the time change in log *Knowledge Access* from 1951 to 1966, averaged across technologies within each MSA. Darker colors represent higher growth in *Knowledge Access*. As with patent growth, we observe that MSAs that had the strongest growth are generally located in the South and the West of the United States, far from the knowledge centers of New York and Chicago. The reduction in travel time was larger between locations far apart, implying that locations which happened to be far from knowledge centers increased relatively more their *Knowledge Access*.

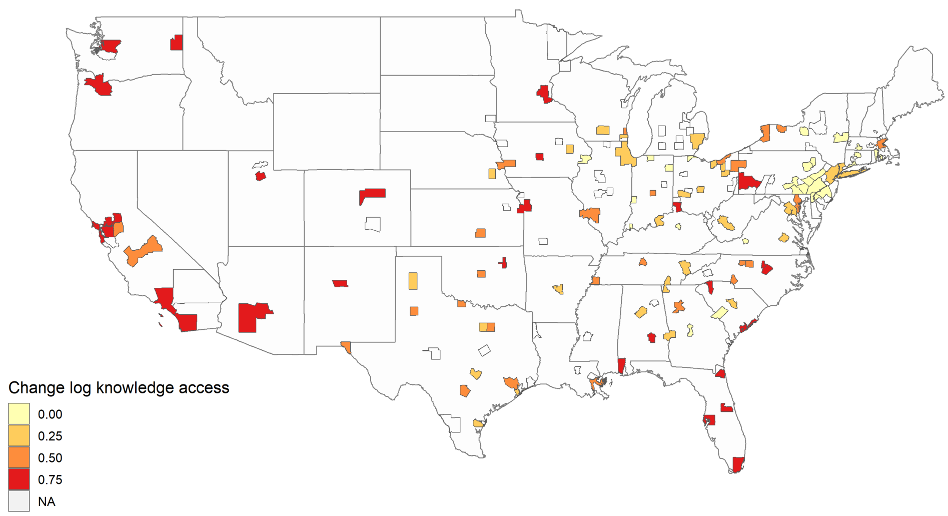


Fig. 1.16. Change in log Knowledge Access 1951 - 1966

With the measure of *Knowledge Access* we then adapt equation (1.1) to estimate:

$$\text{Patents}_{Fiht} \;=\; \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht} \tag{1.7}$$

where Patents_{Fiht} are patents applied by establishment of firm F in location i and technology h at time period t. The measure of knowledge access KA_{iht} is at the iht location-technology-time level, meaning that all establishments within an iht share the same level of knowledge access. The parameter of interest ρ is the elasticity of (the creation of new) patents to knowledge access. In the presence of knowledge spillovers as suggested in section 2, we would expect ρ to be positive and statistically significant.

The fixed effect FE_{Fih} absorbs time invariant characteristics at the firm-location-technology level, as for example the productivity of the establishment-technology. This fixed effect is more fine grained than just a location-technology, which would absorb the comparative advantage of a location in a certain technology. The fixed effect FE_{it} absorbs characteristics that are time variant at the location level. For example, changes in R&D subsidies that are location specific and common across all technologies would be absorbed by this fixed effect. Also, better flight connectivity could spur economic activity as shown in Campante and Yanagizawa-Drott (2017), leading to an increase in patenting activity in the location. If that increase is general across technologies within the location, then FE_{it} would absorb it. Finally, the fixed effect FE_{ht} absorbs characteristics that are time variant at the technology level. If technologies had different time-trends at the national level, then the fixed effect would control for these trends in a flexible way.

The inclusion of FE_{Fih} implies that only across-time variation within an establishment-technology is used to identify ρ. The inclusion of FE_{it} implies that only variation across-technologies within a location-time is exploited, so across-time variation is compared across establishments within a location, and not across locations. The inclusion of FE_{ht} implies that the identifying across-time variation is conditional on aggregate trends of the technology. In short, identification of ρ relies on across-time changes in the amount of patents and knowledge access of an establishment, relative to other establishments in the same location, conditional on aggregate technological trends.

Column (1) in Table 1.4 shows the result of estimating equation ((1.7)). The elasticity of patents to knowledge access is estimated to be 10.14, significant at the one percent level. The average change in knowledge access at the location-technology level[77] is 9%, implying that on average the change in travel time predicts a 3.5% average yearly growth rate of patents.[78] The observed average yearly growth rate of new patents at the location-technology is 4.4%.[79] Comparing the predicted and observed growth rates, the improvement in air travel time has the power to account for 79.5% of the observed average

[77]Due to entry, we cannot compute the growth rate at the establishment-technology level for 70% of establishment-technology, given that they had 0 patents in the initial time period. In the case of location-technology, 5% did not have patents in the initial period. We the prefer to interpret coefficients using location-technology growth rates, which we compute using the remaining 95% of location-technologies that had positive patents in the initial time period.

[78]The elasticity of 10.14 predicts an increase of 91.3% over the time period of 19 years ($10.14 \times 0.09 = 0.913$), which translates into a 3.5% average yearly growth rate ($(1+0.913)^{1/19}-1 \approx 0.035$).

[79]From the first time period (1949-1953) to the last time period (1964-1968) we observe an average growth rate of new patents of 127%. We obtain $0.044 \approx ((1+1.27)^{1/19}-1$

	PPML	PPML q innovation	IV PPML	IV PPML q innovation
Dependent Variable: *Patents*		$Patents_{Fiht}$		
	(1)	(2)	(3)	(4)
log(knowledge access)	10.14^{***} (3.66)	9.36^{**} (3.69)	11.24^{*} (6.35)	10.26 (6.38)
log(knowledge access) × 3rd quartile		2.05^{***} (0.58)		2.32^{***} (0.66)
log(knowledge access) × 2nd quartile		3.80^{***} (0.90)		4.21^{***} (0.84)
log(knowledge access) × 1st quartile		5.00^{***} (1.30)		5.77^{***} (1.11)
R2	0.85	0.85	0.85	0.85
N obs. effective	991,480	991,480	991,480	991,480

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.4: Effect of knowledge access on patents, by MSA innovativeness quartile

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents filed in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Column (3) and (4) show the result of two step instrumental variables estimation, where KA_{iht} is instrumented with $\widetilde{KA}_{iht}$, knowledge access computed using the counterfactual travel time that would have taken place if routes were fixed to the ones in 1951 and each year routes were operated at the average aggregate flying speed of the year. Standard errors are presented in parentheses. Column (1) and (2) present clustered at the location-technology ih. Column (3) and (4) present bootstrap standard errors. R2 is computed as the squared correlation between observed and fitted values.

yearly patent growth rate.[80]

We aggregate predicted changes in patent growth at the Census Region level. The change in travel time predicts a yearly growth rate 0.74 percentage points higher in the South and the West relative to the Midwest and Northeast. In the data we observe 2.1 percentage points difference in the growth rate, implying that the change in travel time can account for 35% of the observed differential growth rate.[81]

Section 5.1 showed that in the data, initially less innovative MSAs had a larger growth rate of patenting. In column (2) in Table 1.4 we investigate if the increase in knowledge access had an heterogeneous effect on the amount of new patents created depending on the initial innovativeness of the location i in technology h. We compute the quartile of innovativeness of location i in technology h in the time period 1949-1953 and interact it with $\log(KA_{iht})$.[82] We use as reference category the highest quartile of initial innovativeness, hence the coefficient on $\log(KA_{iht})$ without interaction is the elasticity for the highest quartile. Coefficients on other quartiles should be interpreted relative to the highest quartile.

We find that the coefficients on lower quartiles of initial innovativeness are positive and statistically different from the coefficient in the highest quartile. Thus, knowledge access had a greater effect on patenting for establishments that were located in initially less innovative locations.[83] Given the difference in the coefficients, the increase in knowledge access predicts an average yearly growth of new patents of 4.5% for the initially lowest quartile of innovativeness, while it predicts 3.4% for the highest quartile.[84] The change

[80] $79.5 = 3.5/4.4 \times 100$

[81] Using the coefficient of column (1) in Table 1.4, we compute the MSA-technology predicted level of patents for 1966 and aggregate it at the Census region - technology level. Then, we compute yearly growth rates within each region-technology and take averages across technologies. Next, we take the average between S and W, and MW and NE, and finally compute the differential predicted growth. If we use the quartile-specific coefficients of column (2) in Table 1.4 we obtain a predicted differential growth rate of 0.86 percentage points, which implies that the change in travel time can account for 41% of the observed differential growth rate.

[82] We use the quartiles of innovativeness defined in section 5.1, computed using the amount of patents of location i in technology h filed in the time period 1949-1953. Each location i has (potentially) a different value quartile in each technology h. The 1st quartile refers to the 25% initially least innovative MSAs in technology h.

[83] A given percentage change in knowledge access led to a stronger increase in patenting in initially less innovative locations.

[84] The change in knowledge access for the lowest quartile is on average 9.1%, which multiplied by the coefficient 14.36 (obtained by doing 9.36+5.00=14.36) gives a predicted growth of 131% over 19 years. Translated into average yearly growth it is 4.5% = $[(1+1.31)^{(1/19)} - 1] \times 100$. For the highest quartile, knowledge access changed on average 9.5%, which multiplied by the coefficient 9.36 predicts 89% growth

in knowledge access predicts differential growth rate of 1.1 percentage points. In the data we observe that the average yearly growth rate of patents in the lowest quartile is 5.3 percentage points higher than in the highest quartile. Comparing the predicted and observed differential growth rates, the improvement in knowledge access as consequence of the reduction in travel time explains 21% of the difference in growth rates of new patents between locations in the lowest and highest quartile of innovativeness.[85]

As in Section 6, we may be concerned that decisions of the regulator or airlines which affect travel time are endogenous to the diffusion of knowledge and consequently to knowledge access. Therefore, we construct an instrument for knowledge access in which instead of using observed travel time, we use the fictitious travel time presented in section 4.2 in which routes are fixed to the ones in 1951 and each route is operated with the average airplane of the year:

$$\widetilde{KA}_{iht} = \sum_{k} \omega_{hk} \sum_{j, j \neq i} \text{Patent stock}_{jk,t=1953} \times \left(\text{travel time}_{ijt}^{\text{fix routes}}\right)^{\beta} \tag{1.8}$$

We then implement the instrumental variables estimation by control function as in Section 6. The results are presented in columns (3) and (4) in Table 1.4. The coefficients do not show an important change and the convergence prediction obtained using non-instrumented PPML remains valid.[86,87]

Figure 1.17 shows in the left panel the patent growth observed in the data (it replicates Figure 1.9), while in the right panel it is the predicted patent growth. We compute the prediction using the observed change in travel time and quartile specific elasticities of column (2) in Table 1.4. Similarly to what is observed in the data, the change in travel time predicts a larger patenting growth rate in the South and the West. At the same time, the change in travel time predicts smaller growth rates in New York, Chicago and their surroundings.

The result in column (2) implies that a given change in *Knowledge Access* had a stronger ef-

rate, which is 3.4% yearly growth rate.

[85] $21\% \approx 1.2/5.1 \times 100$

[86] The first stage of the model estimated in column (3) of Table 1.4 gives a $\hat{\lambda}_2 = 1.01$ with standard error 0.03 (clustered at the location-technology level *ih*), and a within R2 of 0.53.

[87] Using IV estimates, the predicted yearly patent growth rate in the lowest quartile is 4.9% while it is 3.7% in the highest quartile. The predicted differential growth rate is then 1.2 percentage points, meaning that the change in knowledge access can explain $(1.2/5.3) \times 100 \approx 23\%$ of the observed differential growth rate.

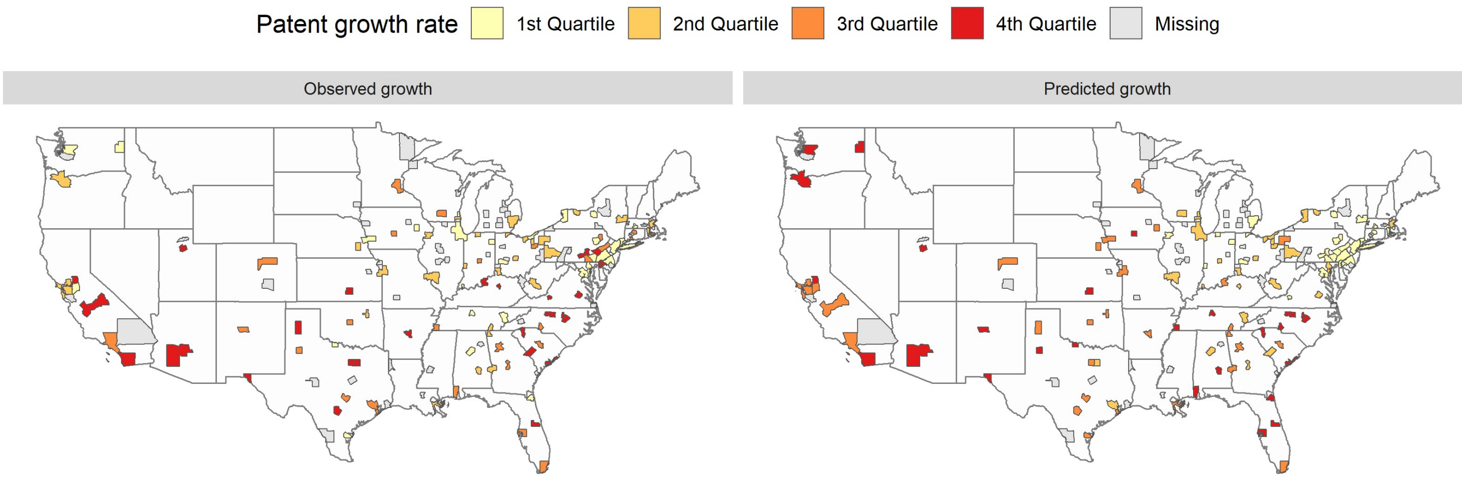


Fig. 1.17. Observed vs. predicted patent growth 1951 - 1966

fect on patenting growth in less innovative locations. In other words, knowledge spillovers as an externality had a more predominant role in the production of knowledge in locations that initially produced relatively fewer patents. Theoretically, this result implies that the parameter ρ in equation ((1.1)) varies depending on the level of previous production of knowledge of location i. Empirically the implication is that a given increase in knowledge spillovers leads to innovation convergence across locations. As seen in section 5.1, during 1949-1968 we observe innovation-convergence across locations and that is exactly what the estimated coefficients predict following a reduction in travel time.

In order to understand the convergence result and compare it with other findings in the literature it is important to remember that commercial airplanes during 1950s and 1960s were a means of transportation mainly for people. On the other hand, other transportation improvements as those in water transport, railroads or highways also contain another ingredient: they were used to carry goods. Hence, other means of transportation have a simultaneous impact on face to face interactions and trade. Pascali (2017) finds that the introduction of the steam engine vessels in the second half of the 19th century had an impact on international trade that led to economic divergence between countries. Faber (2014) finds that the expansion of the highway system in China led to a reduction of GDP growth in peripheral counties, with evidence suggesting a trade channel due to reduction in trade costs. In our setup, the introduction of jet airplanes represented a big shock to the mobility of people while not affecting significantly the transport of merchandise. Therefore, studying the introduction of jet airplanes allows us to focus on improved face

to face interactions, while the trade channel would be a second order effect.

7.1. *Creation of knowledge: Robustness*

In this section we show that the effect of *Knowledge Access* on the creation of new patents and the convergence effect remains after including different controls. Table 1.5 shows the results.

Jaworski and Kitchens (2019) show that improvements in the Interstate Highway System led to local increases in income through an increased market access. In our set up, if the effect of market access affects innovation in the same way across technologies, then it would be absorbed by the MSA-time fixed effect FE_{it} in equation ((1.7)). However, if the effect of market access on innovation varies across technologies, then it would be a confounder. To control for this potential confounder, we compute market access by highway and interact it with a technology dummy. We compute market access as:

$$\text{Market Access}_{it} = \sum_{j} \text{Population}_{j,t=1950} \times \tau_{ijt}^{\theta} \tag{1.9}$$

where $\text{Population}_{j,t=1950}$ is population in MSA j in 1950, τ_{ijt} are the shipping costs provided in the data of Taylor Jaworski and Carl Kitchens computed using each year's highway driving distance, highway travel time, petrol cost and truck driver's wage. θ is the elasticity of trade to trade costs which we set to -8.28, the preferred value of Eaton and Kortum (2002) and in the range of many other estimates in the literature (Head and Mayer (2014), Caliendo and Parro (2015), Donaldson and Hornbeck (2016)). Columns (3) and (4) of Table 1.5 show the results, we do not observe an important difference with the baseline estimates.

Campante and Yanagizawa-Drott (2017) shows that better connectivity by airplane leads to an increase in economic activity as measured by satellite-measured night light. Söderlund (2020) shows that an increase in business travel in the late 1980s and early 1990s led to an increase in trade between countries. In a similar way to market access, we could think that better connectivity by airplane could have led to an increase in market access due to a reduction in information frictions, with goods being shipped by land. Similarly to highway market access, if the effect of market access by airplane is common to all technology categories the effect would be absorbed by the MSA-time fixed effect FE_{it}. In order to account for a technology-specific effect, we construct a measure of airplane market access and interact it with a technology dummy. The measure of airplane market access is similar to equation (1.9) where τ is the travel time by airplane and θ is set to -1,22, the elasticity of

trade to travel time from Söderlund (2020). The results are shown in columns (5) and (6) of Table 1.5. While the coefficients in all quartiles are reduced, the estimated value of ρ is positive and significant and the result on convergence remains.

Potential contemporaneous improvements in other means of communication, like telephones, could have spurred the creation of new patents. In columns (7) and (8) we include the log of the MSA's share of households with telephones in 1960 and double-interact it with a technology dummy and a time dummy. The results remain invariant with respect to the baseline.

Another potential explanation for the increase of patenting could be that better connectivity decreased technology-specific financial frictions. The potential reduction in financial frictions, rather than a confounder, would be a mechanism through which airplanes increased innovation. However, according to Jayaratne and Strahan (1996) during 1950s and 1960s interstate lending and bank branching were limited. Prior to the 1970s, banks and holdings were restricted in their geographic expansion within and across state borders. Additionally, the Douglas Amendment to the Bank Holding Company Act prevented holding companies from acquiring banks in other states. Therefore, it is unlikely that interstate bank financing would be a driving force. Nonetheless, if other sector-specific modes of financing like venture capital were active, it could be confounding the results. In Appendix E we construct multiple measures of access to capital by using market capitalization of patenting firms listed in the stock market. The results present suggestive evidence that access to capital is not driving the results.

Finally, in Appendix E we include additional robustness checks. We compute different versions of *Knowledge Access*: we use distance-specific β from section 6, we consider the patent stock only of locations j far from i, we do sensitivity analysis using different values of β. Also, we re estimate the effects by quartile of initial innovativeness using patents per capita. Last, we re-do the baseline regression using OLS estimation. Results go in the same direction: an increase in knowledge access leads to an increase in patenting and the effect is stronger in initially less innovative locations.

8. Firms' geographic expansion

In section 5.1 we showed that there was innovation-convergence across regions and this happened simultaneously with an increase in the amount of multi-establishment firms.

	PPML							
Dependent Variable: *Patents*	$Patents_{Fiht}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(knowledge access)	10.14***	9.36**	9.28**	8.23**	6.22*	5.84	10.34***	9.25***
	(3.66)	(3.69)	(3.68)	(3.69)	(3.58)	(3.60)	(3.44)	(3.43)
log(knowledge access) × 3rd quartile		2.05***		2.16***		2.06***		2.23***
		(0.58)		(0.57)		(0.59)		(0.57)
log(knowledge access) × 2nd quartile		3.80***		3.89***		3.75***		3.93***
		(0.90)		(0.89)		(0.88)		(0.91)
log(knowledge access) × 1st quartile		5.00***		5.13***		5.08***		5.18***
		(1.30)		(1.30)		(1.29)		(1.32)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
Controls:								
log(Highway market access) × technology	-	-	Yes	Yes	-	-	-	-
log(Airplane market access) × technology	-	-	-	-	Yes	Yes	-	-
log(Telephone share) × technology × time	-	-	-	-	-	-	Yes	Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.5: Elasticity of new patents to knowledge access, by MSA innovativeness quartile

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of Patents$_{Fiht}$ $=$ exp $[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Relative to columns (1) and (2), columns (3) and (4) control for technology specific effect of log(highway market access), columns (5) and (6) control for technology specific effect of log(airplane market access), columns (7) and (8) control for technology and time specific effect of log(telephone share). Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

In section 7 we showed that the reduction in travel time predicts innovation-convergence across locations. In this section we uncover one of the mechanisms that led to innovation-convergence: the geographic expansion of multi-establishment firms. We proceed in two steps. First, we show that the increase in patenting is driven by two types of entry: entry of establishments of new firms, and entry of establishments of pre-existing firms. The second type of entry is due to the geographic expansion of firms. Second, we show that the decrease in travel time led firms to expand geographically and this expansion was stronger towards initially less innovative locations.

8.1. Entry of new establishments

We use all patents of the same firm to identify all locations in which the firm had research establishments in each time period.[88] Using patents applied during the first time period (1949-1953), we classify all the research establishments that applied for patents in every subsequent period. We classify research establishments into three mutually exclusive categories: the establishment (and hence the firm) applied for patents in 1949-1953 (*existing firm and est*), the establishment did not apply for patents but the firm had establishments in other locations applying for patents in 1949-1953 (*existing firm new est*), neither the establishment nor the firm applied for patents in 1949-1953 (*new firm new est*).[89] The dummies *new firm new est* and *existing firm new est* capture two types of entry margin. *new firm new est* captures a new establishment of a new firm, while *existing firm new est* captures entry due to the geographic expansion of firms. The dummy *existing firm and est* captures jointly an intensive and exit margin.

We estimate a variation of equation ((1.7)) that includes interactions with dummies which indicate the status of the establishment in 1949-1953:

$$\text{Patents}_{Fiht} = \exp\left[\sum_{e} \boldsymbol{\rho_e}\, log(KA_{iht}) \times \mathbb{1}\{Fi \in e\} + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \nu_{Fiht} \quad (1.10)$$

where Patents_{Fiht} are patents applied by establishment of firm F in location i and technology h at time period t. KA_{iht} is the knowledge access at the location-technology-time level. $\mathbb{1}\{Fi \in e\}$ is an indicator variable that takes value 1 of Fi is of the type $e = \{$*new firm new*

[88] All our *firm* and *research establishment* information comes from the patent data. Hence, we only observe an establishment in a certain time period if it applies for patents in that time period.

[89] We define if an establishment exists or not if it applied for patents in any technology h in 1949-1953. We define the establishment at the Fi level (as opposed to Fih) as our object of interest a firm-location. An interesting avenue of research is to study within-establishment changes in the technological composition of patenting.

est, *existing firm new est*, *existing firm and est*$\}$. The results are displayed in column (2) of Table 1.6. The results show that the effect of innovation access on the increase of patenting happened through the two entry margins: entry of new establishments of new firms and entry of new establishments of firms that previously existed in other locations.

Dependent Variable: *Patents*	$Patents_{Fiht}$ (1)	(2)
log(knowledge access)	10.14^{***} (3.66)	
log(knowledge access) $\times$ new firm new est		23.71^{***} (4.46)
log(knowledge access) $\times$ existing firm new est		23.79^{***} (4.47)
log(knowledge access) $\times$ existing firm and est		-0.28 (4.70)
R2	0.85	0.81
N obs. effective	991,480	991,480

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.6: Patents and knowledge access: Entry, exit and continuing firms
Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \zeta_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) adds an interaction of $log(KA_{iht})$ with e the type of establishment Fi in a classification on whether the establishment and/or the firm existed in 1949-1953. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

In Table 1.7 we open up the effect by including a double interaction of Fi establishment type and location-technology ih quartile of initial innovativeness. We use the highest quartile as the reference category. The two margins of entry are active in all quartiles of initial innovativeness, with a stronger effect in lower quartiles. In the case of the entry of establishments that belong to firms that already existed in other locations, the pattern is more prominent. The intensive and exit margin does not appear active in any quartile of innovativeness except for the last one. The combined effect of entry and intensive/exit suggests that, in locations in the lowest quartile of initial innovativeness, the churn rate of patenting firms is increased as consequence of the increase in knowledge access.

The results of Table 1.6 and Table 1.7 indicate that one part of the increase in patenting is consequence of multi-establishment firms that expand across locations, and more so in initially less innovative locations. Hence, multi-establishment firms contributed to innovation-convergence across locations by expanding geographically.

Quartile innovativeness \ Establishment type	New firm & New est	Existing firm & New est	Existing firm & Existing est
log(knowledge access)	22.84^{***}	22.00^{***}	-0.36
	(4.40)	(4.41)	(4.67)
log(knowledge access) × 3rd quartile	3.40^{***}	6.35^{***}	-1.33
	(1.14)	(1.44)	(1.19)
log(knowledge access) × 2nd quartile	5.95^{***}	6.74^{***}	-2.20
	(1.48)	(1.67)	(2.33)
log(knowledge access) × 1st quartile	4.88^{**}	10.98^{***}	-15.62^{***}
	(1.97)	(2.15)	(3.25)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.7: Patents and knowledge access: entry, exit and continuing firms

The table shows the results of one Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\sum_e \boldsymbol{\rho_e}\, log(KA_{iht}) \times \mathbb{1}\{Fi \in e\} + \sum_{q\in\{1,2,3\},e} \boldsymbol{\rho_{q,e}}\, log(KA_{iht}) \times \mathbb{1}\{ih \in q\} \times \mathbb{1}\{Fi \in e\} + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \nu_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{ikt} is knowledge access of establishments in location i technology h and time period t. q is the quartile of initial innovativeness of location i within technology h, computed using patents filed in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. e is the type of establishment Fi in a classification on whether the establishment and/or the firm existed in 1949-1953. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values. All columns and rows belong to the same regression. The number of observations is 991,480.

8.2. Geographic expansion of multi-establishment firms

In this subsection we show that the decrease in travel time gave rise to the geographic expansion of multi-establishment firms. We focus on all firms that patented in the initial time period and follow their subsequent opening and closure of establishments. We find that firms directed the opening (closure) of new establishments towards locations that got larger (smaller) reductions in travel time to the firm's headquarters.

We define the headquarters location q of firm F as the location in which the firm filed the largest amount of patents in the period 1945-1953. If firm F did not file any patent in 1945-1953, or there is no unique location with the maximum amount of patents (e.g. two locations have the maximum amount of patents), then no headquarters is assigned.[90] Firms with no headquarters assigned are dropped from the estimations that required

[90] Using patents applied in the period 1949-1953 does not significantly affect the results. We use 1945-1953 instead as it allows us to identify headquarters location for 7% more firms.

headquarters location.

We compute the travel time of every firm F's headquarters's location q to each other location j. We then estimate a linear probability model to study if the location decision of establishments of a firm depend on travel time to a firm's headquarters. We estimate the following regression:

$$\mathbb{1}\{establishment_{Fqjt}\} = \gamma\, log(\text{travel time}_{qjt}) + FE_{Fqj} + FE_{Fqt} + FE_{jt} + \zeta_{Fqjt} \tag{1.11}$$

where $\mathbb{1}\{establishment_{Fqjt}\}$ is a dummy variable that takes value 1 if firm F with headquarters in location q has a research establishment in location j at time period t.[91] The coefficient γ is a semi-elasticity: $\gamma/100$ is the change in percentage points of the probability that firm F has an establishment in location j when travel time increases by one percent. If travel time has a negative impact on the probability then we would expect γ to be negative.

The inclusion of the fixed effect FE_{Fqj} implies that γ is identified only from changes in travel time and *opening* and *closure* of research establishments across time.[92] Fixed effects FE_{Fqt} and FE_{jt} control flexibly for changes in firm F expanding and the opening establishments everywhere else, and j becoming more attractive for every firm.

Table 1.8 presents the results jointly with predicted and observed growth rate of the probability. Column (1) presents the results of estimating equation (1.11). We find that the probability of firm F having a subsidiary research establishment in location j increases when the travel time between the firm's headquarters's location q and j decreases. The coefficient is -0.0364, which if we multiply it by the average change in travel time between headquarters' location and every other potential location (-34.7%), the decrease in travel time predicts an increase in the share of existing subsidiaries of 0.0126 percentage points. The result goes in the same direction as Giroud (2013) who finds that a reduction in travel time between a firm's subsidiary and its headquarters leads to an increase in investment in the subsidiary.

Column (2) of Table 1.8 estimates the semi-elasticity of the probability of having an establishment to travel time by the quartile of innovativeness of location j in 1949-1953.

[91] $\mathbb{1}\{establishment_{Fqjt}\}$ takes value 0 if firm F does not file patents in location j at time period t. The headquarters location q remains fixed for all time periods.

[92] *Opening* refers from $\mathbb{1}\{establishment_{Fqjt}\}$ switching from 0 to 1, while *closure* refers to the inverse.

	Baseline	Quartile receiving location	Initial probability	Change travel time	Predicted yearly growth rate	Observed yearly growth rate
Dependent Variable:	$\mathbb{1}\{establishment_{Fqjt}\}$					
	(1)	(2)				
log(travel time)	-0.0364*** (0.0088)		0.000810	-34.7%	15.94%	1.50%
log(travel time) × 4th quartile		-0.0749*** (0.0187)	0.001895	-36.0%	15.41%	0.98 %
log(travel time) × 3rd quartile		-0.0150*** (0.0031)	0.000364	-33.4%	15.22%	3.03%
log(travel time) × 2nd quartile		-0.0102*** (0.0028)	0.000145	-35.2%	18.67%	3.86%
log(travel time) × 1st quartile		-0.0079*** (0.0025)	0.000068	-33.8%	21.40%	5.75%
R2	0.49	0.50				
N obs. effective	19,755,792	19,755,792				

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.8: Subsidiaries' location and travel time to headquarters

The table shows the estimation of a linear probability model. The left panel of the table shows estimation results while the right panel shows observed and predicted growth rates of the probability. Column (1) presents the results of OLS estimation of $\mathbb{1}\{establishment_{Fqjt}\} = \gamma\, log(\text{travel time}_{qjt}) + FE_{Fqj} + FE_{Fqt} + FE_{jt} + \zeta_{Fqjt}$ or firm F which has headquarters in location $q$where $\mathbb{1}\{establishment_{Fqjt}\}$ is a dummy that takes value one if firm F which has headquarters in location q has an establishment *open* in location j at time period t. We define an establishment of firm F in location j at time period t as *open* if F has inventors located in j that apply for patents at time period t. travel time$_{qjt}$ is the travel time in minutes between F's headquarters location q and location j at time period t. Column (2) includes an interaction of $log(\text{travel time}_{qjt})$ with the across-technology average quartile of initial level of innovativeness of j. j's quantile of initial innovativeness in technology h is computed using the level of patents of j in 1949-1953 in technology h. Standard errors at the non-directional location pair are presented in parentheses (qj is the same non-directional location pair as jq). Predicted growth rates are obtained using the estimated coefficient and the change in travel time, relative to the initial probability. Yearly growth rates g are obtained by computing $g = [(1 + \text{nineteen year growth rate})^{(1/19)} - 1] \times 100$, where 19 is the amount of years between 1949 and 1968.

We compute the quartile of innovativeness at the location level by taking the average quantile across technologies within a location, only for those technologies in which the location has positive patents in 1949-1953. To be able to compare the relative impact of travel time on the growth rate of the probability, we need to relate the semi-elasticity to the baseline probability. The semi-elasticity in the lowest quartile of initial innovativeness is around 1/10th the one in the highest quartile. However, the initial probability in the lowest quartile is around 1/30th of the one in the highest quartile. Therefore, a given percentage change in travel time has an impact on the growth rate of the probability in the lowest quartile that is around 3 times the one in the highest quartile.[93] In other words, given the very low initial probability of locations in the lowest quartile of innovativeness to receive a subsidiary from a firm headquartered in another location, the small increase in percentage points represents a big relative increase in the probability.

The yearly growth rate of subsidiaries implied by the change in travel time is 21.4% for the lowest quartile while it is 15.4% for the highest quartile, implying a predicted difference of 6 percentage points in the yearly growth rate.[94] In the data we observe an average yearly growth rate which is 4.8 percentage points higher for the lowest quartile relative to the highest quartile.[95] Hence, the reduction in travel time not only predicts a geographic expansion of firms, but it also predicts that the geographic expansion is tilted towards initially less innovative locations. This pattern of geographic expansion is in line with the one observed in the data.

[93] These are approximate numbers. The precise computations: the ratio of coefficients is $0.106 = (-0.0079)/(-0.0749)$, the ratio of initial probability is $0.036 = 0.000068/0.001895$, the ratio of the growth rate is $2.94 = (-0.0079/0.000068)/(-0.0749/0.001895)$. The initial probabilities are computed as the amount of observed subsidiaries in 1949-1953 divided by the amount of (time invariant) potential subsidiaries. The amount of potential subsidiaries is the amount of firms for which we identify headquarters multiplied by the amount of locations other than headquarters location (we have 108 locations in the data, meaning that each firm has 107 potential locations for subsidiaries).

[94] For the lowest quartile, the model predicts a 3,869% increase in the probability over 19 years ($19 = 1968 - 1949$), which translates into an average yearly growth rate of 21.4%. For the highest quartile the predicted increase is 1,422%, an average yearly growth rate of 15.4%. Consistent with the computation of the relative growth rate presented in the main text: $1,422/3,869 = 0.36 \approx 0.34 \times (33.8/36.0)$, where 0.34 has to be adjusted by the fact that the average change in travel time is not the same across quartiles. The 19-year growth rates are obtained by multiplying the change in travel time (-33.8% vs -36.0%) by the coefficient (-0.0079 vs -0.0749) divided by 100, and finally dividing by the initial probability (0.000069 vs 0.001895) and multiplying by 100. For the lowest quartile: $3,869 = [(-33.8) \times (-0.0079/100)/0.000069] \times 100$, and for the highest quartile:$[1,422 = (-36.0) \times (-0.0749/100)/0.001895] \times 100$. The average yearly growth rates are computed as $21.4 \approx [(1+38.69)^{1/19} - 1] \times 100$ and $15.4 \approx [(1+14.22)^{1/19} - 1] \times 100$.

[95] The average yearly growth rate of the probability for the lowest quartile is 5.75% while it is 0.98% for the highest quartile.

9. Conclusion

This paper constructed a new dataset of the flight network in the United States during the beginning of the *Jet Age* and studied the impact of improvements of air travel on the creation and diffusion of knowledge. We found that the reduction in travel time led to an increase in knowledge diffusion, especially between research establishments located far apart. The reduction in travel time also led to an increase in the general access to knowledge, which had positive spillovers for the creation of new knowledge. The effect in the increase of creation of knowledge was stronger in locations initially less innovative, generating a convergence force which goes in the same direction as what is observed in the data. One of the drivers of the increase in the creation of knowledge and convergence was the geographical expansion of firms.

We provide causal evidence of *standing on the shoulders of giants*: new knowledge builds upon pre-existing knowledge. We do so by first estimating one new key parameter: the elasticity of diffusion of knowledge to travel time. Second, extending a production function of knowledge proposed in Carlino and Kerr (2015), we estimate the impact of knowledge spillovers on the creation of new knowledge. Conditional on the pre-existing distribution of knowledge, changes in travel time translate into changes in knowledge spillovers. The results show that knowledge spillovers are important for the creation of new knowledge and more so in locations which are initially less innovative.

Our novel dataset document a historical country wide event that dramatically changed the way we see time and space. Our results provide new evidence of how the introduction of jet airplanes changed the geography of innovation. Better connectivity to innovation centers in the Midwest and the Northeast led to an increase in innovation in the South and the West of the United States. In this way, jet airplanes were one of the contributing factors in the shift of innovative activity towards the South and the West of the United States.

We would like to point to the limitations of the current analysis. The results found in this paper are identified by exploiting differential time changes across establishments. As consequence, we are able to identify differential impacts and not aggregate ones. The results obtained could be consequence of a general increase in the amount of diffusion and creation of knowledge, a relocation of previous diffusion and creation, or a mix of both. At the same time, the potential relocation of resources resulting from the reduction in travel time may have increased the allocative efficiency and thereby the amount of knowledge produced.

References

Acemoglu, D., Akcigit, U., and Kerr, W. R. (2016). Innovation network. *Proceedings of the National Academy of Sciences*, 113(41):11483–11488.

Aghion, P. and Howitt, P. (1997). *Endogenous Growth Theory*. The MIT Press.

Agrawal, A., Galasso, A., and Oettl, A. (2017). Roads and innovation. *The Review of Economics and Statistics*, 99(3):417–434.

Andersson, D., Berger, T., and Prawitz, E. (2017). On the right track: Railroads, mobility and innovation during two centuries'. In *On the Move: Essays on the Economic and Political Development of Sweden*, pages 203–288. Institute for International Economic Studies, Stockholm University.

Andrews, M. J. and Whalley, A. (2021). 150 years of the geography of innovation. *Regional Science and Urban Economics*, page 103627.

Arzaghi, M. and Henderson, J. V. (2008). Networking off madison avenue. *The Review of Economic Studies*, 75(4):1011–1038.

Audretsch, D. B. and Feldman, M. P. (2004). Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics*, volume 4, pages 2713–2739. Elsevier.

Bai, J. J., Jin, W., and Zhou, S. (2021). Proximity and knowledge spillovers: Evidence from the introduction of new airline routes.

Balsmeier, B., Assaf, M., Chesebro, T., Fierro, G., Johnson, K., Johnson, S., Li, G.-C., Lück, S., O'Reagan, D., Yeh, B., Zang, G., and Fleming, L. (2018). Machine learning and natural language processing on the patent corpus: Data, tools, and new measures. *Journal of Economics & Management Strategy*, 27(3):535–553.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The quarterly journal of economics*, 122(2):775–805.

Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers*, (13).

Bloom, N., Schankerman, M., and Van Reenen, J. (2013). Identifying technology spillovers and product market rivalry. *Econometrica*, 81(4):1347–1393.

Borenstein, S. and Rose, N. L. (2014). *How Airline Markets Work... or Do They? Regulatory Reform in the Airline Industry*, pages 63–135. University of Chicago Press.

C.A.B. (1951, 1956, 1961, 1966). *Air carrier traffic statistics.* Civil Aeronautics Board.

Caliendo, L. and Parro, F. (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies*, 82(1):1–44.

Campante, F. and Yanagizawa-Drott, D. (2017). Long-Range Growth: Economic Development in the Global Network of Air Links*. *The Quarterly Journal of Economics*, 133(3):1395–1458.

Carlino, G. and Kerr, W. R. (2015). Agglomeration and innovation. *Handbook of regional and urban economics*, 5:349–404.

Catalini, C., Fons-Rosen, C., and Gaulé, P. (2020). How do travel costs shape collaboration? *Management Science*, 66(8):3340–3360.

Caves, R. E. (1962). *Air transport and its regulators: an industry study*. Harvard economic studies v. 120. Harvard University Press, Cambridge.

Coscia, M., Neffke, F., and Hausmann, R. (2020). Knowledge diffusion in the network of international business travel. *Nature Human Behaviour*, 4(10).

De Rassenfosse, G. and Jaffe, A. B. (2017). Econometric evidence on the r&d depreciation rate. Technical report, National Bureau of Economic Research.

Dhaene, G. and Jochmans, K. (2015). Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies*, 82(3):991–1030.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Duranton, G., Martin, P., Mayer, T., and Mayneris, F. (2009). The economics of clusters: evidence from france.

Duranton, G. and Puga, D. (2004). Micro-foundations of urban agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2063–2117. Elsevier.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from china's national trunk highway system. *Review of Economic Studies*, 81(3):1046–1070.

Fogel, R. W. (1963). *Railroads and American economic growth: essays in econometric history*. PhD thesis, Johns Hopkins University.

Furman, J. L. and Stern, S. (2011). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review*, 101(5):1933–63.

Giroud, X. (2013). Proximity and Investment: Evidence from Plant-Level Data *. *The Quarterly Journal of Economics*, 128(2):861–915.

Glaeser, E. (2011). *Triumph of the City*. Macmillan.

Gross, D. P. (2019). The consequences of invention secrecy: Evidence from the uspto patent secrecy program in world war ii. Technical report, National Bureau of Economic Research.

Haines, M. R. et al. (2010). Historical, demographic, economic, and social data: the united states, 1790–2002. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*.

Hall, B. H., Jaffe, A. B., and Trajtenberg, M. (2001). The nber patent citation data file: Lessons, insights and methodological tools. Working Paper 8498, National Bureau of Economic Research.

Hansen, B. E. (2021). *Econometrics*.

Head, K. and Mayer, T. (2014). Chapter 3 - gravity equations: Workhorse,toolkit, and cookbook. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, pages 131–195. Elsevier.

Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics*, 122:103316.

Hovhannisyan, N. and Keller, W. (2015). International business travel: an engine of innovation? *Journal of Economic Growth*, 20(1):75–104.

Jaffe, A. B., Trajtenberg, M., and Fogarty, M. S. (2000). Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review*, 90(2):215–218.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Jaworski, T. and Kitchens, C. T. (2019). National policy for regional development: Historical evidence from appalachian highways. *Review of Economics and Statistics*, 101(5):777–790.

Jayaratne, J. and Strahan, P. E. (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics*, 111(3):639–670.

Jones, C. I. (2002). Sources of u.s. economic growth in a world of ideas. *American Economic Review*, 92(1):220–239.

Kerr, W. R. and Robert-Nicoud, F. (2020). Tech clusters. *Journal of Economic Perspectives*, 34(3):50–76.

Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological Innovation, Resource Allocation, and Growth*. *The Quarterly Journal of Economics*, 132(2):665–712.

Krugman, P. R. (1991). *Geography and trade*. MIT press.

Lucas, R. E. (1993). Making a miracle. *Econometrica*, 61(2):251–272.

Magerman, T., Van Looy, B., and Song, X. (2006). Data production methods for harmonized patent statistics : patentee name harmonization.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., and Ruggles, S. (2020). Ipums national historical geographic information system: Version 15.0.

Marshall, A. (1890). *Principles of economics*.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics*, 90(4):683–701.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, 111(10):3328–75.

Pascali, L. (2017). The wind of change: Maritime technology, trade, and economic development. *American Economic Review*, 107(9):2821–54.

Perlman, E. R. (2016). Dense enough to be brilliant: patents, urbanization, and transportation in nineteenth century america. *Work. Pap., Boston Univ*.

Petralia, S. (2019). HistPat International Dataset.

Petralia, S., Balland, P.-A., and Rigby, D. (2016). HistPat Dataset.

Redding, S. and Venables, A. J. (2004). Economic geography and international inequality. *Journal of international Economics*, 62(1):53–82.

Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of econometrics*, 190(2):212–221.

Silva, J. M. C. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Söderlund, B. (2020). The importance of business travel for trade: Evidence from the liberalization of the soviet airspace. Working Paper Series 1355, Research Institute of Industrial Economics.

Staiger, D. and Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65(3):557–586.

Stock, J. H. and Yogo, M. (2005). *Testing for Weak Instruments in Linear IV Regression*, page 80–108. Cambridge University Press.

Storper, M. and Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4(4):351–370.

Tsiachtsiras, G. (2021). Transportation networks and the rise of the knowledge economy in 19th century france.

Weidner, M. and Zylkin, T. (2021). Bias and consistency in three-way gravity models. *Journal of International Economics*, 132:103513.

Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics*, 182(1):226–234.

Appendix

A. Travel Time Data

A.1. Data Construction

We construct a dataset of travel times by plane between US MSAs for the years 1951, 1956, 1961, 1966. We get information of direct flights from airline flight schedules and feed this information into an algorithm to allow for indirect flights. For each MSA pair with airports served by at least one of the airlines in our dataset we compute the fastest travel time in each of the four years.

Using images of flight schedules, we digitized the flight network for six major airlines: American Airlines (AA), Eastern Air Lines (EA), Trans World Airlines (TWA), United Airlines (UA), Braniff International Airways (BN) and Northwest Airlines (NW). Note that the first four in this list were often referred to as the *Big Four*, highlighting their dominant position in the market. They alone accounted for 74% of domestic trunk revenue passenger-miles from February 1955 to January 1956. Together the six airlines accounted for 82% of revenue passenger-miles in that same period, 77% from February 1960 to January 1961 and 78% from February 1965 to January 1966 (C.A.B., 1966). Our sample of airlines thus covers a vast share of the domestic market for air transport. In addition, the airlines were chosen to maximize geographic coverage.

In total we obtain a sample of 5,910 flights. These flights often have multiple stops. If we count each origin-destination pair of these flights separately, our sample contains 17,469 legs.

Table 1.9 lists the exact dates of when flight schedules we digitized became effective. Due to limited data availability not all flight schedules are drawn from the same part of the year. As seasonality of the network seems limited and given the large market share of the airlines we consider, our data is a good approximation of the network in a given year.

Table 1.9: Date of Digitized Flight Schedules

Airline	1951	1956	1961	1966
AA	September 30	April 29	April 30	April 24
EA	August 1	October 28	April 1	April 24
TWA	August 1	September 1	April 30	May 23
UA	April 29	July 1	June 1	April 24
BN	August	August 15	April 30	April 24
NW	April 29	April 29	May 28	March 1
PA	June 1	July 1	August 1	August 1

Figure 1.18 shows two pages of the flight schedule published by American Airlines in 1961. Each column corresponds to one flight. As can be seen, one flight often has multiple stops. Departure and arrival times in most flight schedules are indicated using the 12-hour system. PM times can be distinguished from AM times by their bold print. In the process of digitization we converted the flight schedules to the 24-hour system. Times in most tables are in local time. We thus recorded the time zones that are indicated next to the city name and converted them to Eastern Standard Time.

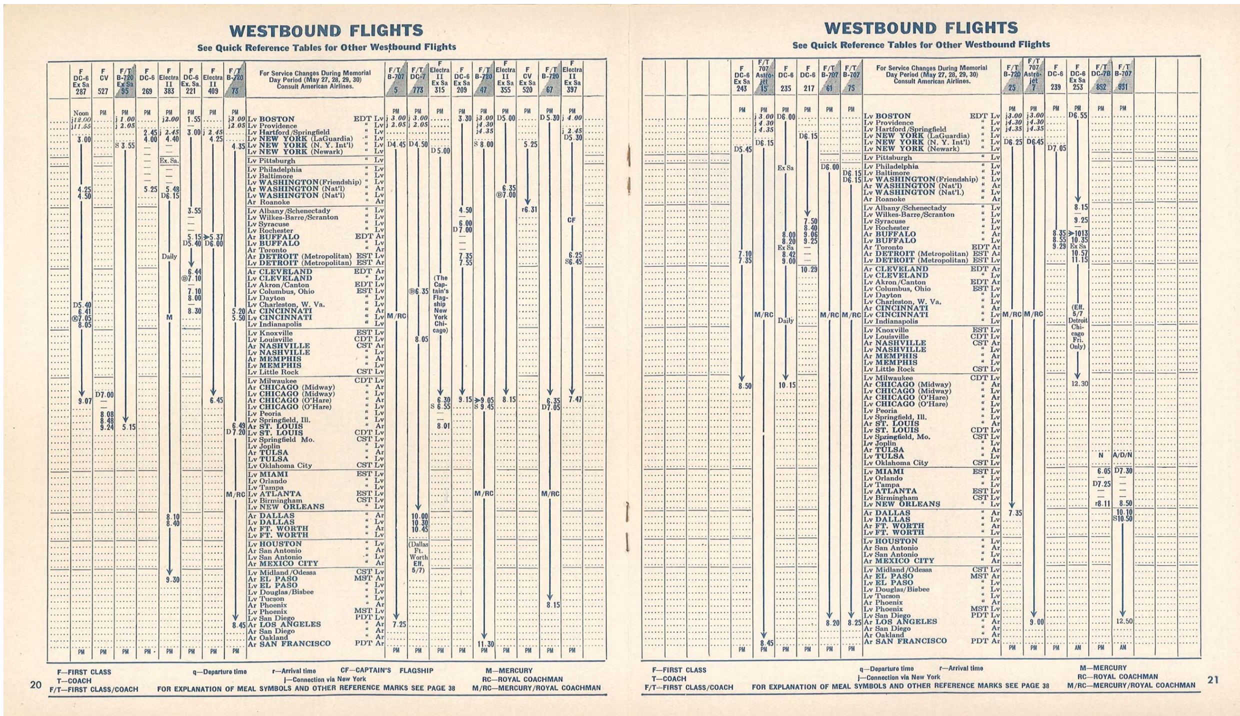


Fig. 1.18. Flight Schedule American Airlines 1961.

To obtain exact geographical information on where airports are located, we match city names to their IATA airport codes. We use the addresses of ticket offices that are indicated on the last pages of the flight schedules. Most of the ticket offices were located directly at the airport, allowing to infer the airport the airline was serving in a given year. For some flight schedules we are missing these last pages and used information from adjacent years in order to identify airports. We also manually verified the airport match using various online sources. We then obtain geographical coordinates from a dataset provided by `https://ourairports.com/` (downloaded July 2020).

From the flight schedule we also collect information on the aircraft model, indicated next to the flight number. Using various online sources, we manually identified aircraft models that are powered by a jet engine. We thus know on which connections airlines were using jet aircraft.

Flight Schedules also contain information on connecting flights. For example, the second column in figure 1.18 indicates a departure from Boston leaving at 12.00 local time. A footnote is added to the departure time indicating that this departure is a connection via New York. It is thus not operated by flight 287 otherwise described in column 2, but it is

just supplementary information for the passenger. As we are interested in the speed of aircraft and the actual travel time on a given link, this information on connecting flights would pollute our data and we thus delete this supplementary information.

As outlined above, the digitization requires human input. It is thus prone error-prone. The travel time calculation relies on each link in the network, and if one important connection has a miscoded flight, it might potentially distort the travel time between many MSA pairs. We thus implement an elaborate method to detect mistakes in the digitization process. In particular, after the initial transcription, we regress the observed duration of the flight on a set of explanatory variables: the full interaction of distance, a set of airline indicators, a set of year indicators and a dummy variable indicating whether the aircraft is powered by a jet engine or not. This linear model yields an R^2 above 95%. We then compute the predicted duration of each flight and obtain the relative deviation from the observed duration. If the deviation is above 50%, we manually check whether the transcribed information is correct. If we find a mistake, we correct the raw data, rerun the regression and recompute relative deviations, until all the observations with more than 50% deviation have been manually verified.

For 15 connections, the information was correctly transcribed from the flight schedule, but the flight time differed a lot from other flights with similar distances that used the same aircraft. The implied aircraft speed for these cases is either unrealistically high or low, in one case the implied flight time is even negative. These cases seem to be typos introduced when the flight schedule was created (e.g. a ”2” becomes a ”3”). Instead of inferring what the true flight schedule was which is not always obvious, we drop these cases. Table 1.10 lists all 15 cases.

Table 1.10: Dropped Connections

	Airline	Year	Origin	Destination	Departure Time	Arrival Time
0	UA	66	TYS	DCA	1940	2036
1	UA	66	LAX	BWI	2150	1715
2	UA	66	CHA	TYS	1635	1909
3	PA	66	SFO	LAX	2105	1850
4	PA	66	SEA	PDX	705	935
5	PA	56	PAP	SDQ	830	835
6	PA	51	HAV	MIA	800	903
7	PA	51	SJU	SDQ	825	830
8	NW	66	HND	OKA	655	1135
9	EA	66	ORD	MSP	2340	2340
10	EA	56	SDF	MDW	1352	1418
11	EA	56	GSO	RIC	2207	2204
12	AA	56	PHX	TUS	1630	1655
13	PA	51	STR	FRA	1320	1540
14	EA	66	TPA	JFK	1330	1548

As our analysis is at the MSA level, we match airports to 1950 MSA boundaries. Each airport is matched to all MSAs for which it lies inside the MSA boundary or at most 15km away from the MSA boundary. If we focus only on airports contained within MSA boundaries, we would, for example, drop Atlanta's airport. Of 275 US airports, 156 airports are matched to at least one MSA. 18 of these are matched to two MSAs and Harrisburg International Airport is matched to three MSAs: Harrisburg, Lancaster and York. Out of 168 MSAs, 142 are at some point connected to the flight network in our dataset. In table 1.11 we present the 168 MSAs, the ones that are connected at least once, and the ones that are connected in the four years.

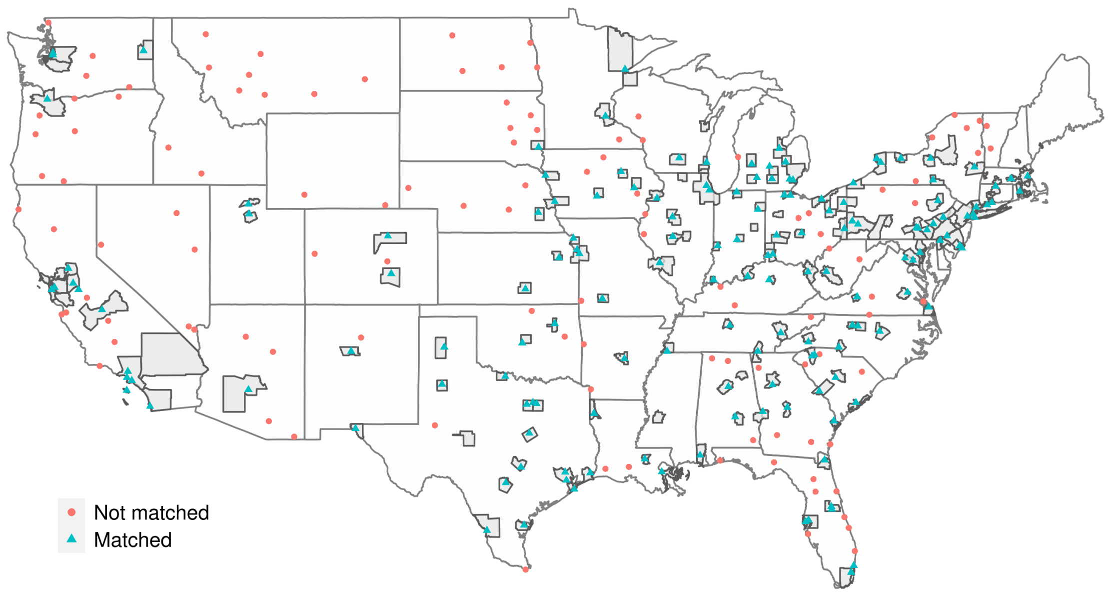


Fig. 1.19. Airports matched to MSAs.

Next, we compute the shortest travel time for every airport pair, and then take the minimum to obtain shortest travel time at the MSA pair level. In particular, we apply Dijkstra's algorithm to compute shortest paths (**?**). We adjust this algorithm to take into account the exact timing of the flight schedules. We consider a possible departure time t from origin city o and then compute the shortest path to destination city d at this time of the day. If getting to d requires switching flights, we account for the required time at the location of the layover. We repeat this procedure for every possible departure time t at origin city o and then take the minimum that gives us the fastest travel time from o to d, τ_{od}.

The flight schedule format requires us to make one assumption. In particular, the flight schedule for a multi-stop flight may either indicate the arrival time or the departure time for a particular stop. If the flight schedule only lists the departure time, we need to infer the arrival time and vice versa. We allow for five minutes between arrival and departure. This is relatively low, but still in the range of observed difference between departure and arrival for cases where we observe both. As correspondences may have been ensured by airlines in reality, i.e. one aircraft waiting with departure until other aircraft arrive, we

opted for the lower end of the observed range of stopping times.

Finally, since the shortest travel time measure may not capture the benefits of a highly frequented hub, we also calculate the daily average of the shortest travel time. In particular, we compute the shortest travel time at every full hour of the day and take the average. This measure thus captures the benefits of being located near an airport where flights depart many times per day.

To conclude, we end up with a set of four origin-destination matrices indicating the fastest travel time (and another set with the average daily travel time) between US MSAs in 1951, 1956, 1961 and 1966.

A.2. Descriptive Statistics

Table 1.12 shows the number of non-stop connections between MSAs by year and airline. It underlines the dominant position of the *Big Four* (AA, EA, TW, UA) which were much bigger than their competitors (BN and NW). The growth of the airline industry is also apparent. All airlines had the lowest number of connections in 1951 and subsequently extended their network. At the same time the average distance of the connections gradually increased over time. Part of this may have been due to jet technology allowing for longer aircraft range. We thus analyze a period where more and longer flights are introduced.

Table 1.12: Domestic Non-Stop Connections by Airline and Year

Airline	Year	Number of connections	Jet Share (connections)	Jet Share (km)	Mean Distance (in km)
AA	1951	258	0.00	0.00	515.32
AA	1956	367	0.00	0.00	889.66
AA	1961	325	22.15	50.50	768.24
AA	1966	282	73.40	89.52	1020.36
BN	1951	96	0.00	0.00	317.90
BN	1956	210	0.00	0.00	380.60
BN	1961	176	8.52	18.84	460.41
BN	1966	150	72.00	76.64	553.09
EA	1951	345	0.00	0.00	319.87
EA	1956	479	0.00	0.00	412.60
EA	1961	595	3.70	13.28	441.42
EA	1966	492	54.47	75.46	569.01
NW	1951	77	0.00	0.00	521.70
NW	1956	95	0.00	0.00	724.77
NW	1961	127	11.02	32.43	824.59
NW	1966	136	77.94	90.86	945.81
TW	1951	210	0.00	0.00	503.69
TW	1956	253	0.00	0.00	711.78
TW	1961	240	28.75	54.63	807.72
TW	1966	265	86.42	96.05	1143.30
UA	1951	291	0.00	0.00	492.88
UA	1956	361	0.00	0.00	714.39
UA	1961	323	31.89	65.32	803.49
UA	1966	533	49.91	79.54	781.38

While these changes in the network are remarkable, airlines were constrained by the regulator in opening new routes. Accordingly, table 1.13 shows that the network remains relatively stable over time with more than three quarters of connections remaining intact

within a five-year window. Interestingly, during the beginning of the jet age (i.e. 1956 to 1961), the network appears to have been especially stable, with only 11% of connections either disappearing or newly being added. Thus, the rise of jet aircraft did not lead to a vast reshaping of the network. Given the very different technology, this may be surprising, but may partly be due to heavy regulation.

The table also shows that newly introduced routes were over long distances whereas those discontinued were operating on shorter distances. When changes in the network took place, they thus seemed to improve the network for places further apart.

Table 1.13: Network Changes (weighted by frequency)

Period	Remain connected	Newly connected	Disconnected
Share of Non-stop Connections (%)			
1951 to 1956	78.47	16.79	4.74
1956 to 1961	88.96	6.43	4.6
1961 to 1966	80.64	12.37	6.99
Mean distance (km)			
1951 to 1956	411	1075	337
1956 to 1961	524	914	972
1961 to 1966	568	769	450

Table 1.14: Network Changes

Period	Remain connected	Newly connected	Disconnected
Connected MSAs			
1951 to 1956	119	7	8
1956 to 1961	122	0	4
1961 to 1966	114	7	8
Non-stop Connections			
1951 to 1956	721	357	124
1956 to 1961	908	231	170
1961 to 1966	912	331	227

Changes in the number of connected MSAs and connections among them. A MSA is connected if in our data it appears as having at least one incoming and one outgoing flight. A non-stop connection refers to a pair of origin MSA-destination MSA between which a non-stop flight operates.

Figure 1.20 shows all non-stop connections in our data weighted by the (log) frequency. Initially, the network was concentrated in the Eastern states and transcontinental routes were not yet established, due to technological limitations. In contrast, in the 1960s, after the jet is introduced, intercontinental routes quickly emerge and are operated at a high frequency. Similarly, direct connections from the Northeast to Florida intensify. The figure echos the findings from table 1.14 which illustrates that the overall number of MSA pairs with a direct connection increases over time.

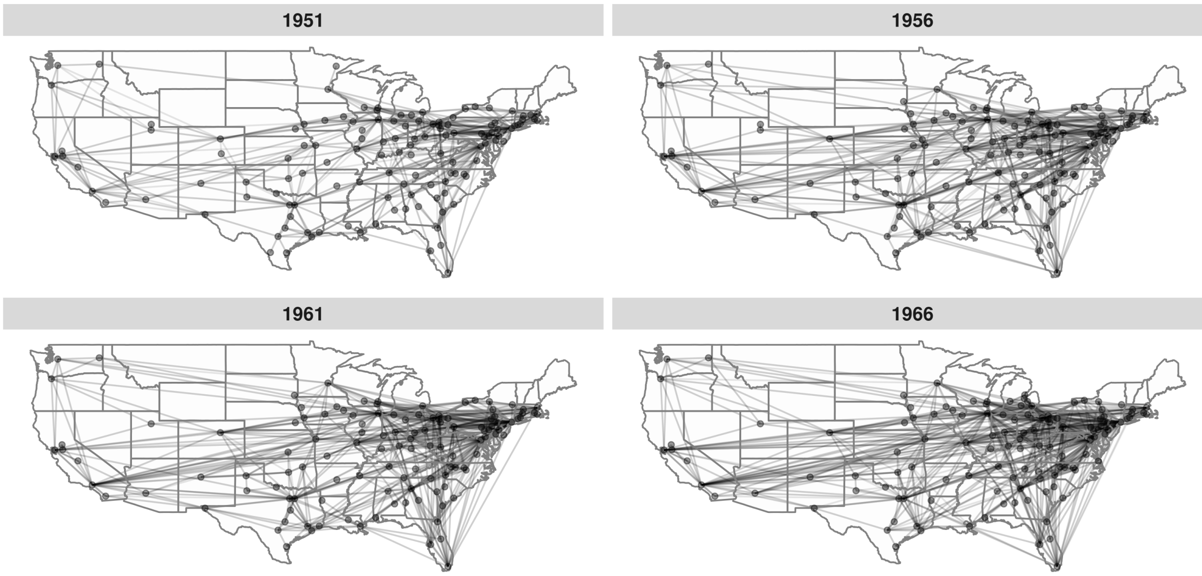


Fig. 1.20. Flight Network by Year. Weighted by log weekly frequency.

Airlines differed in their speed of adoption of the newly arrived jet aircraft. Table 1.12 shows that, in 1961, 65% of UA's connections between MSAs were flown using a jet aircraft (weighted by distance), whereas this was only true for 13% of EA's connections. While adoption was heterogeneous across airlines, adoption was fast. By 1966, all airlines were operating 75% of their connections with jet aircraft (weighted by distance).

Figure 1.21 show the average speed of jet and propeller aircraft by distance. Generally, jet aircraft were substantially faster, but especially so on long-distance flights, where they could be up to twice as fast as propeller-driven aircraft. This particularly stark difference in speed for long-haul flights is also reflected by adoption. Figure 1.22 shows that jet aircraft were first introduced on long-haul flights. Only 50% of MSA pairs at around 1,500 km distance had at least one jet aircraft operating, whereas 100% of pairs above 3,000 km. Then, in the late 1960s, they were also gradually introduced on shorter distances. In fact, for all pairs above 2,000 km there was at least one jet engine-powered flight.

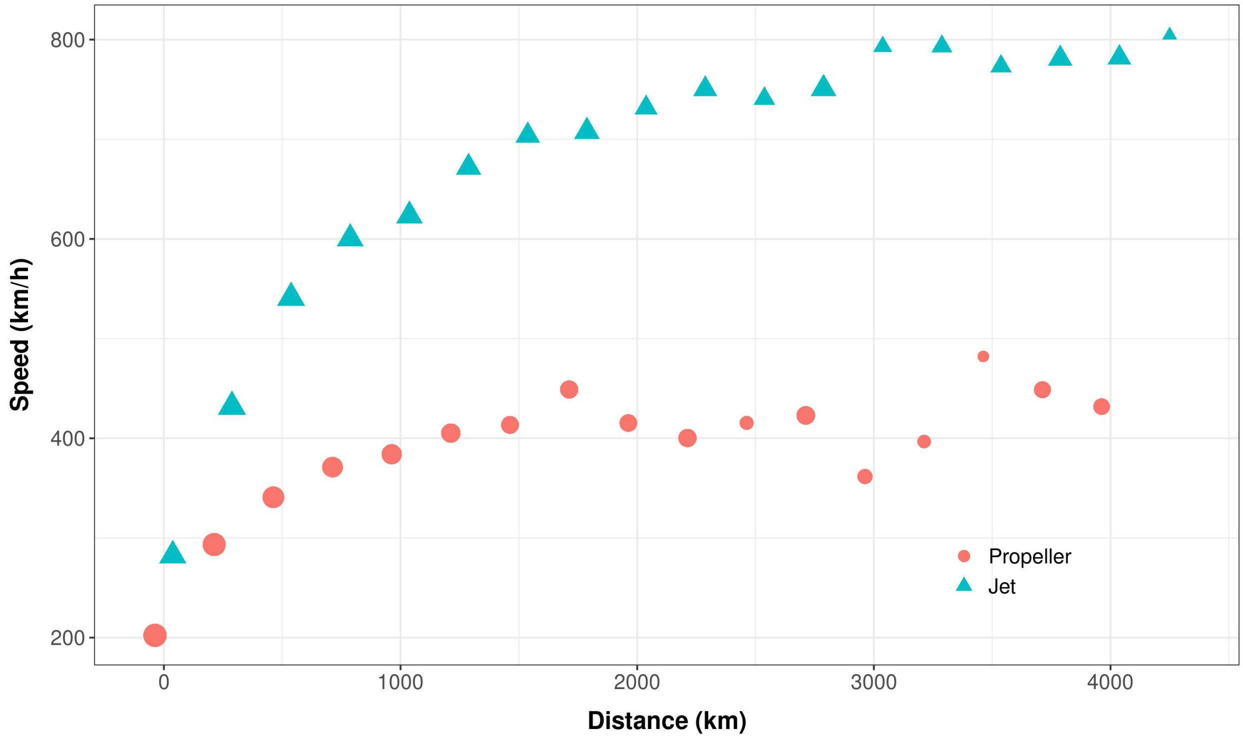


Fig. 1.21. Speed by Aircraft Type. Pooling all Years.

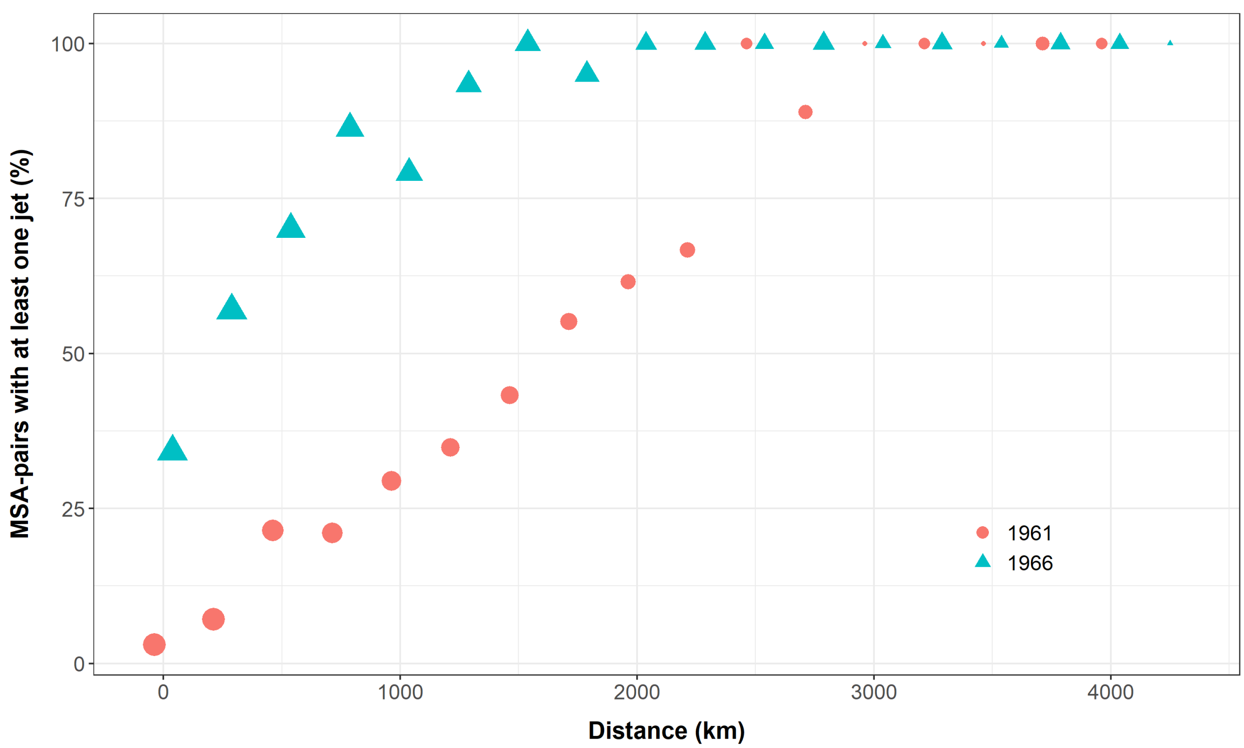


Fig. 1.22. Jet Adoption

Figure 1.23 shows on which routes jets were operating. In the early days of the jet age it was mainly the transcontinental corridor between New York and California that benefited. In 1966 propeller aircraft were already being phased out and only operating in the dense Eastern part of the US where distances between cities are relatively small.

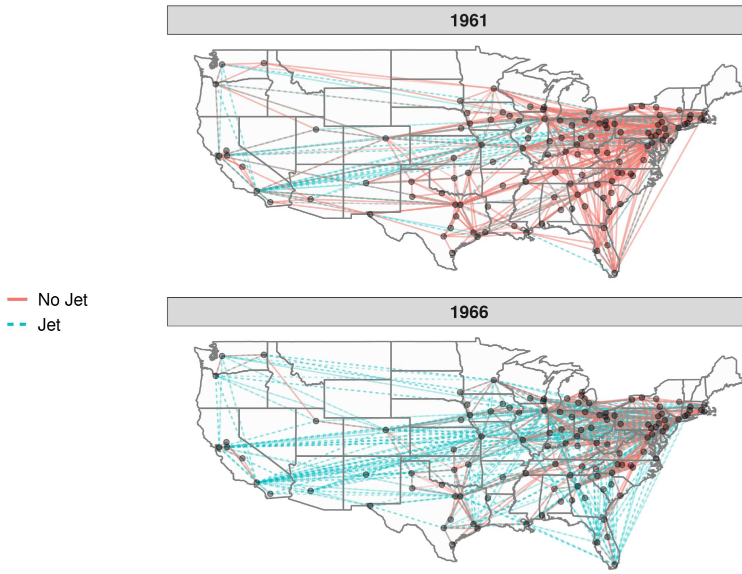


Fig. 1.23. Jet Adoption by Year

The increase in speed due to jet aircraft caused a dramatic reduction in travel times between US cities. When looking at the full origin-destination matrix, i.e. including indirect flights, a network-wide reduction in travel time becomes apparent. Figure 1.24 shows travel times between US MSAs. While the figure shows a gradual decline in travel time from 1951 to 1966, it also illustrates that conditional on distance and year a large amount of variation in travel time remains, as only a small fraction of all MSA pairs were connected via a direct flight (around 8.5% in 1966).

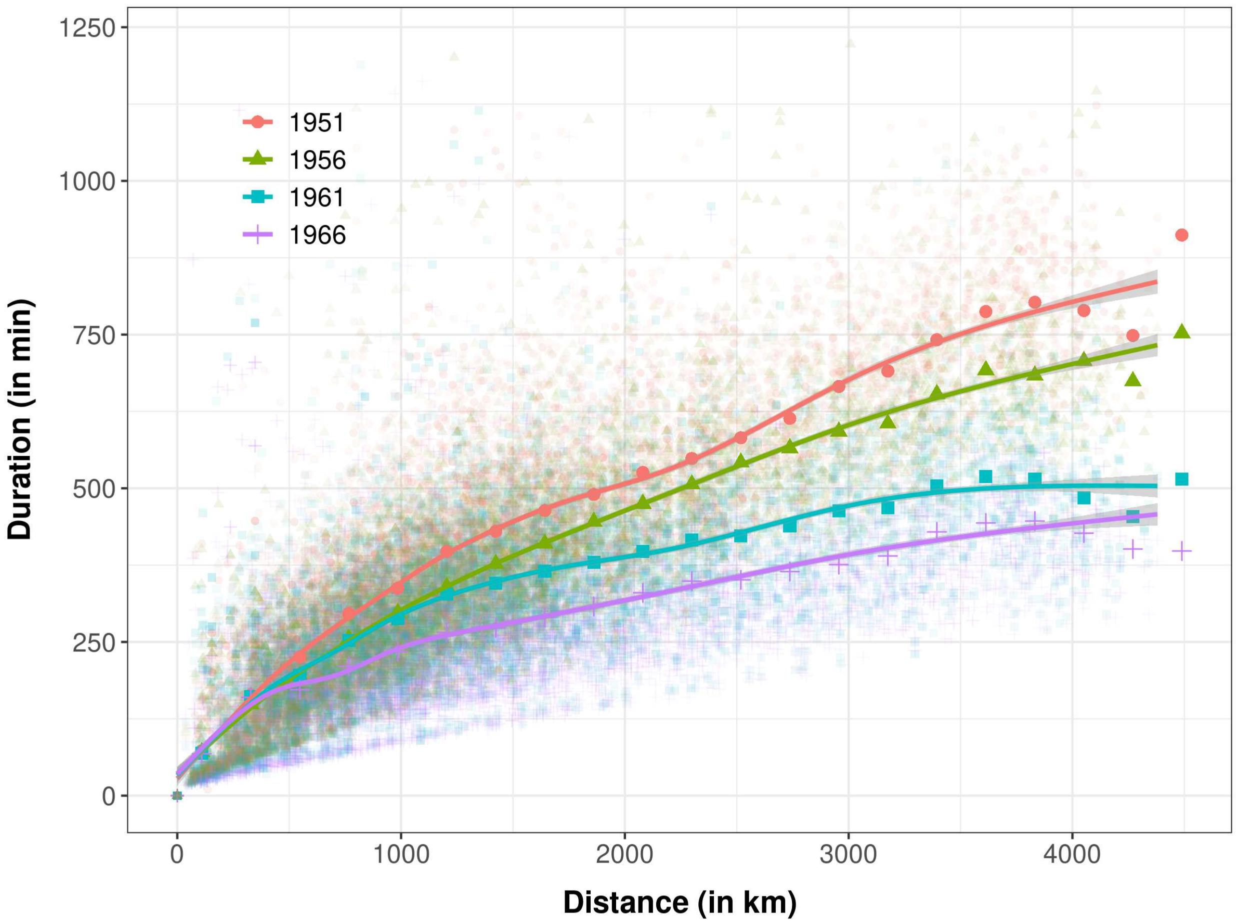


Fig. 1.24. Travel Times between US MSAs.

Figure 1.25 that the change in travel time is accompanied by a reduction of the amount of legs needed to connect two MSAs at every distance. This reduction is specially marked between 1951 and 1956, and 1961 and 1966. In Figure 1.26 we open up the change in travel time by the way an MSA pair was connected in 1951 and 1966: either directly (non-stop flight) or indirectly (connecting flight). We observe that much of the increase in travel time for MSA pairs less than 250km apart comes from routes that were operated non-stop and then it needed a connecting flight. Interestingly, for MSA-pairs more than 2,000km apart travel time reduced on average 42% for those pairs that were connected indirectly in both periods, and 51% for those that switched from indirect to direct. This fact shows the relevance of improvements in flight technology even for MSAs not directly connected. It could be the case that a reduction in the amount of legs or an increase in frequency of flights reduces layover time. In Figure 1.28 we compare the change in travel time from 1951 to 1966 with a fictitious change in travel time in which we eliminate layover time in

both time periods. We observe that the average change in travel time is stronger at every distance if we disregard layover time. This implies that the relative importance of layover time over total travel time increases between 1951 and 1966, preventing total travel time to decrease proportionally to the change of in-flight travel time.

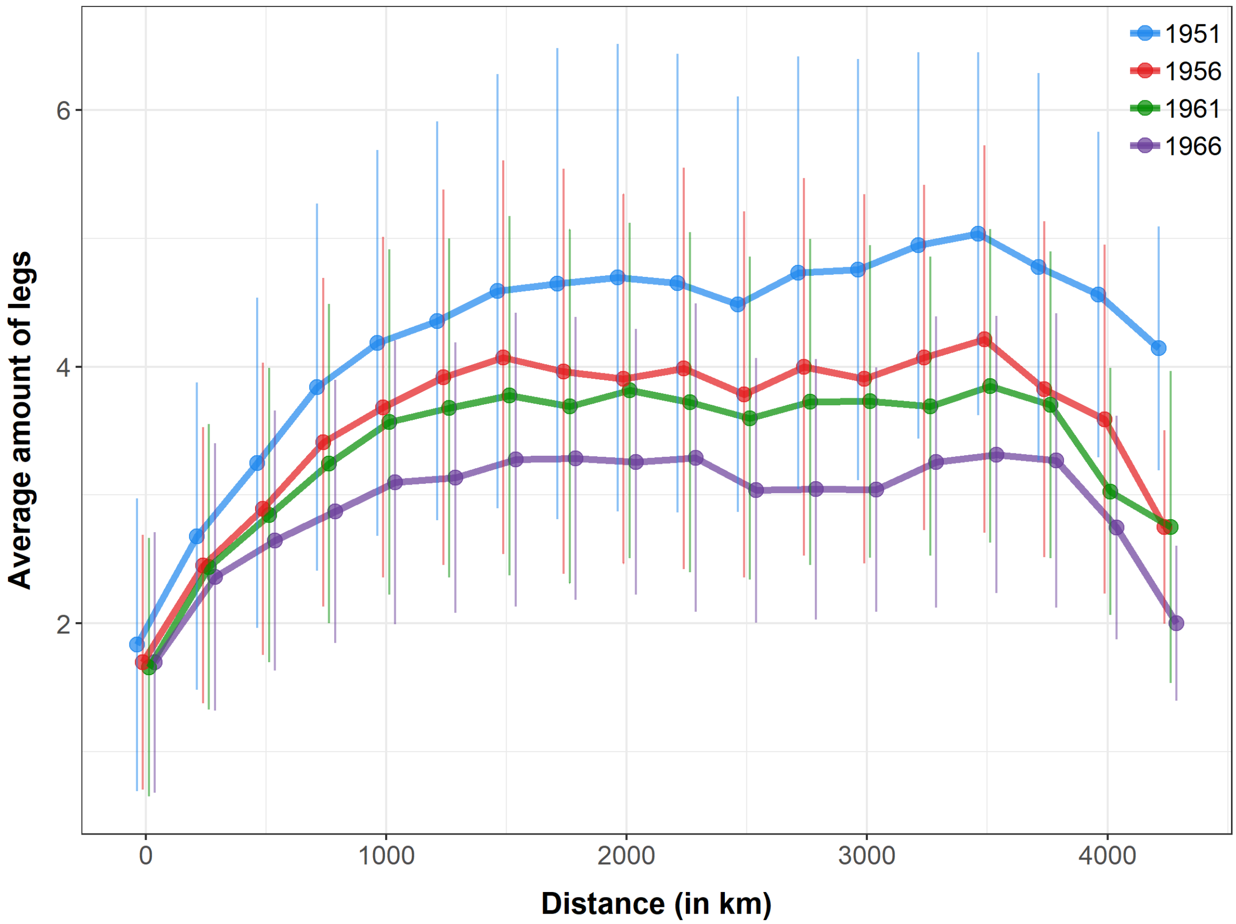


Fig. 1.25. Average amount of legs per route

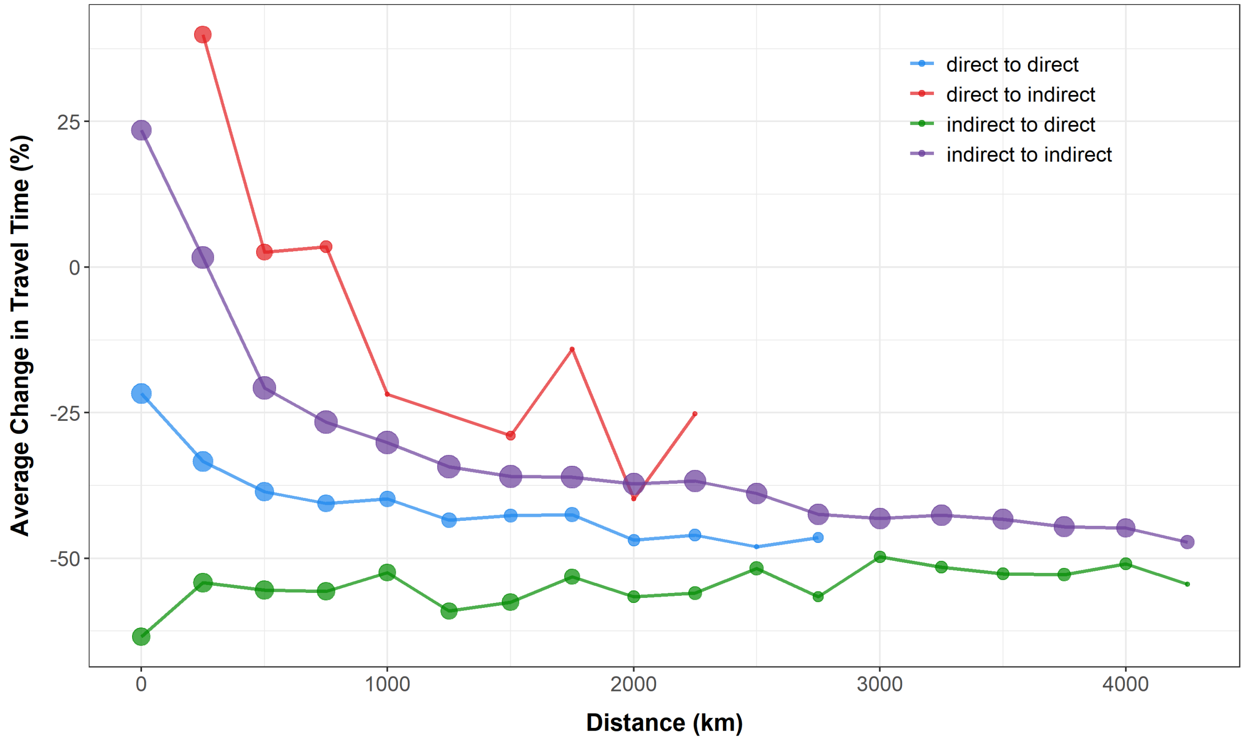


Fig. 1.26. Change in US travel time 1951 to 1966: connections

96

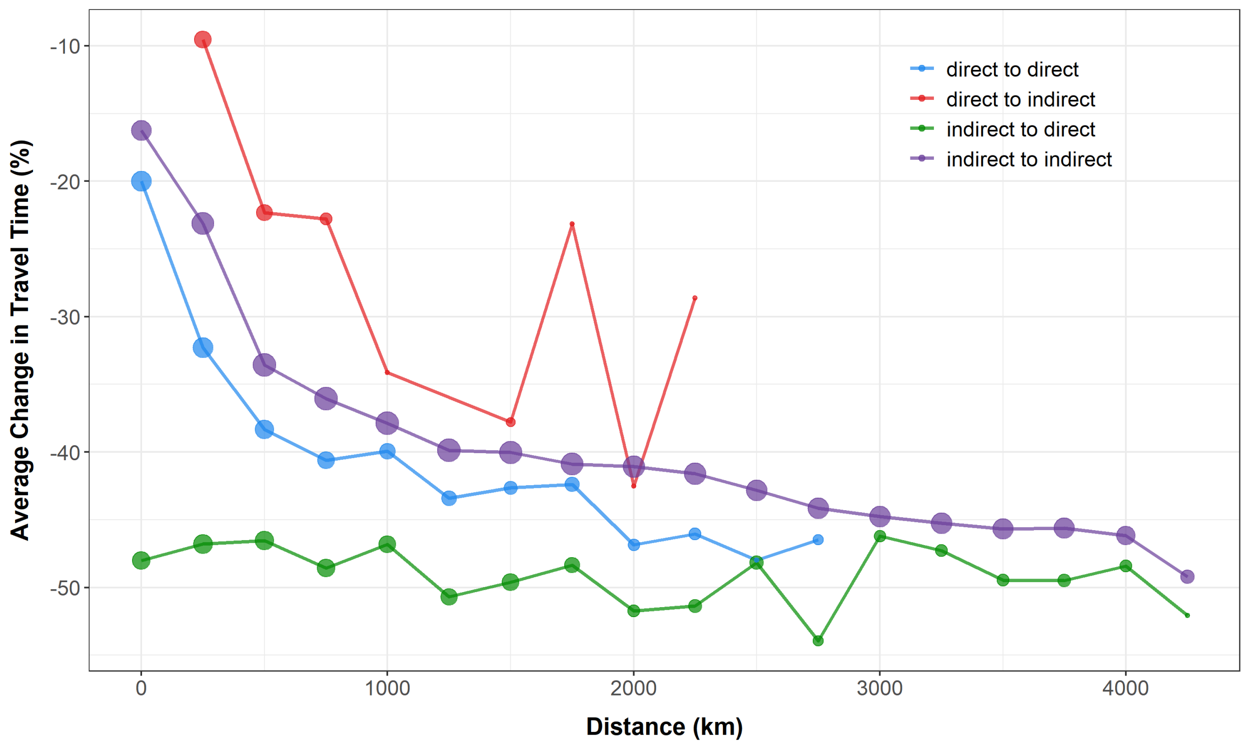


Fig. 1.27. Change in US travel time 1951 to 1966: connections, discarding layover time
97

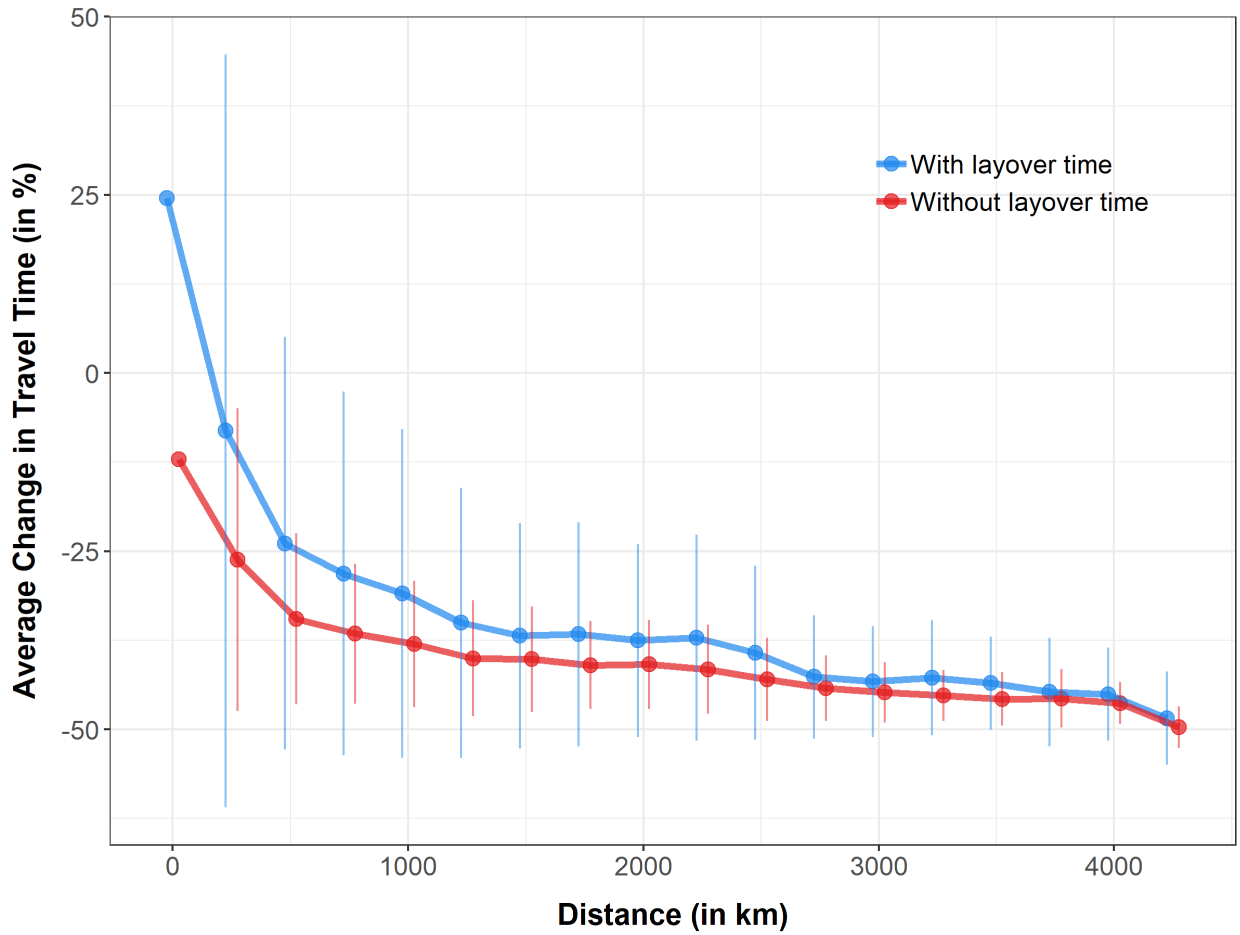


Fig. 1.28. Change in US travel time 1951 to 1966: layover time

In figure 1.29 we show the average change in travel time in three counterfactual flight networks. The first counterfactual fixes the flight routes[98] and allows aircraft speed to evolve. The second counterfactual fixes aircraft speed and allows flight routes to evolve. The third counterfactual allows both flight routes and aircraft speed to evolve. We obtain that around 90% of the change in travel time is due to the change in speed of aircrafts, while around 10% of the change is due to the change in the flight routes. In the figure 1.30 in the appendix we show that the proportion is relatively constant for all distances. This confirms that most of the observed changes in the network are due to improvements in the flight technology.

[98]Fixes the origin-destination airports that are connected with a non-stop flight

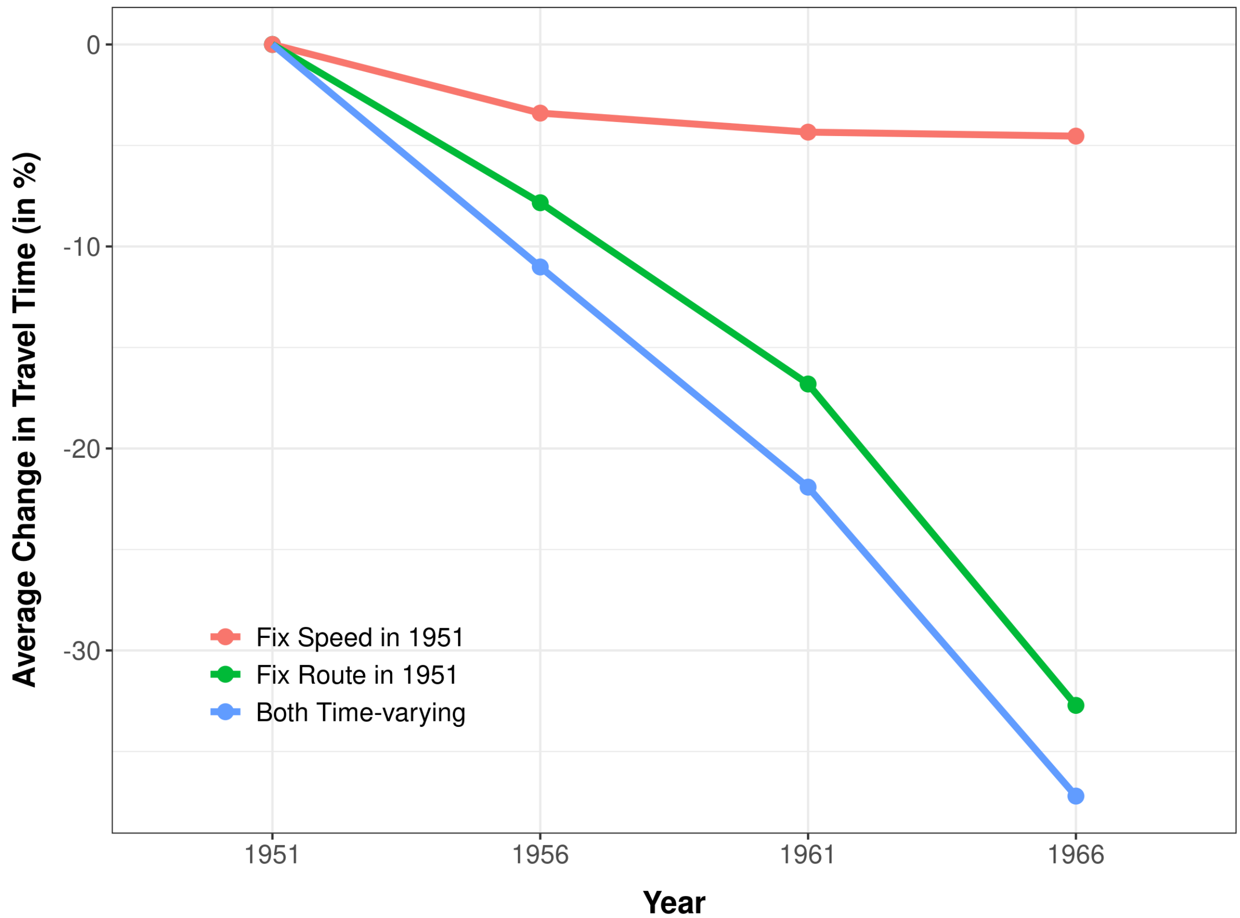


Fig. 1.29. Counterfactual change in travel time

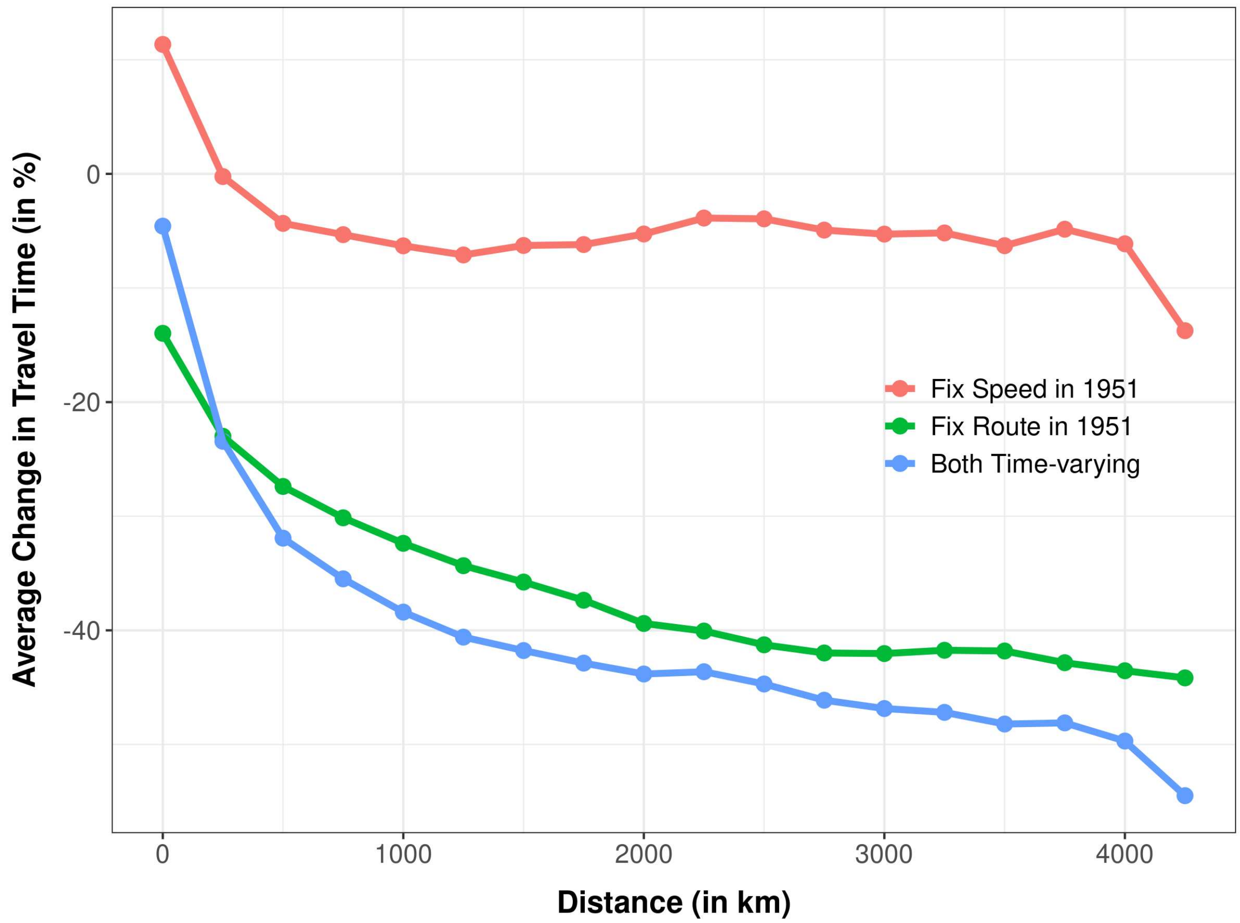


Fig. 1.30. Counterfactual change in travel time 1951-1966

In addition to the changes over time in the network leading to faster travel times, another feature of the US airline industry becomes salient in the data: airlines' regional specialization. As figure 1.31 shows, while there was competition among the airlines in our dataset on the major routes (Lower West Coast to the Midwest and Upper East Coast to the Midwest), some airlines are very specialized and face no competition from any of the other five airlines on certain routes. In particular, NW controls the routes connecting Seattle to the Midwest and EA controls much of the connections from Florida to New York and surroundings.

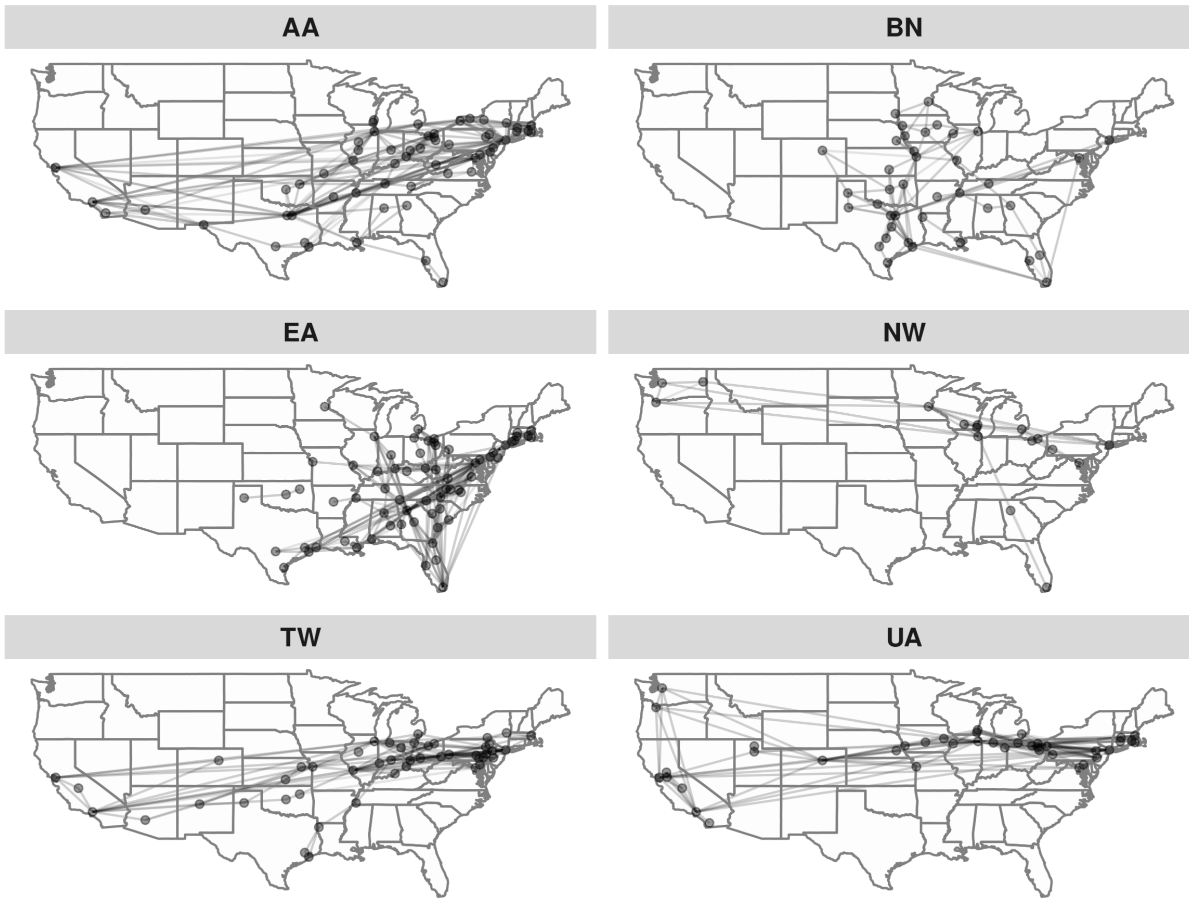


Fig. 1.31. Flight Network in 1956 by Airline (weighted by log frequency).

B. Patent data

In this appendix we describe facts that we observe in the US patent data, for patents filed[99] between 1945 and 1975. US patents data containing citations and filing year have

[99]Filing year, also called application year, is the closest date to the date of invention that is present in the data and it represent the date of the first administrative event in order to obtain a patent. In the other hand, publishing or also called granting year, is the later year in which the patent is granted. The difference between filing and granting year depends on diverse non-innovation related factors (as capacity of the

been downloaded from Google Patents. Then, it was merged with multiple datasets (see Appendix Patent Data Construction for more details):

- Technology classification: NBER patent database.
- Geographic location of inventors: Histpat and Histpat International for patents published until 1975, Fung Institute for patents published after 1975. Both matched to 1950s Metropolitan Statistical Areas (MSAs).
- Ownership: Kogan et al. (2017) for patents owned by firms listed in the US stock market, Patstat for the remaining patents not matched to Kogan et al. (2017).

We highlight two details from the matching process: 1. During filing years 1971-1972 the rate of non-geocoded patents increases, possibly due to Histpat and Fung data not being a perfect continuation one of the other. 2. Kogan et al. (2017) seems to use a matching method based on the patent owner declared in the patent text, as Patstat does. Specially, Kogan et al. (2017) does not explicitly say if it takes into account firm-ownership structure to determine patent ownership, neither does Patstat.

For the analysis presented in this appendix we will use the resulting dataset from the matching procedure, where unless evident or noticed, we will use only patents that have inventors within MSAs. We discard patents that have inventors in multiple MSAs and patents that belong to government organizations or universities. We assign patents to technology categories using fractional count: if a patent is listed in two technology categories, then we assign half a patent to each category. We discard self citations (citations in which the citing patent owner is the same as the cited patent owner) because self-citations may be due to different incentives.

B.1. Matching patents to locations

In figure 1.32 we observe that the matching rate decreases from around 95% before 1970, to around 80% in 1971 and 1972, and then it stabilizes around 99% after 1975. Hence, geogprahical results during years 1970-1975 will contain an increased amount of measurement error.

---

patent office to revise applications) and changes over time. Hence filing year is the date in our data that approximates the best to the date of invention.

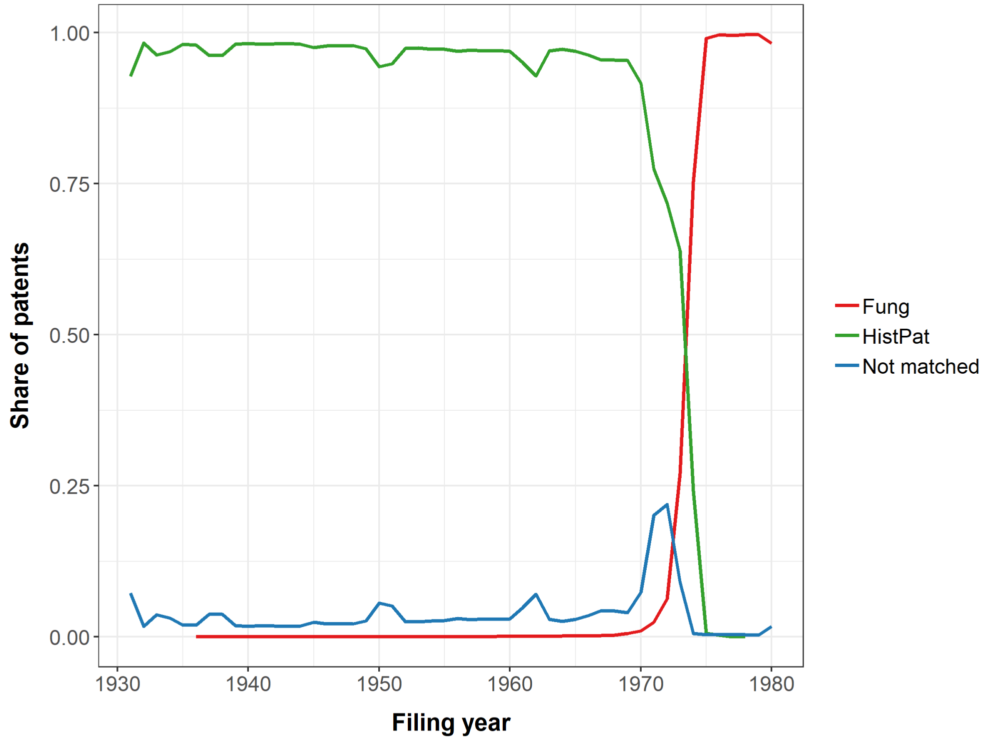


Fig. 1.32. Non-matching rate HistPat, HistPat International and Fung

Figure 1.33 shows the share of patents that have inventors inside MSAs, and figure 1.34 displays the same by technology category.[100]

[100]Technologies are aggregated to six big groups, as explained in HJT 2002

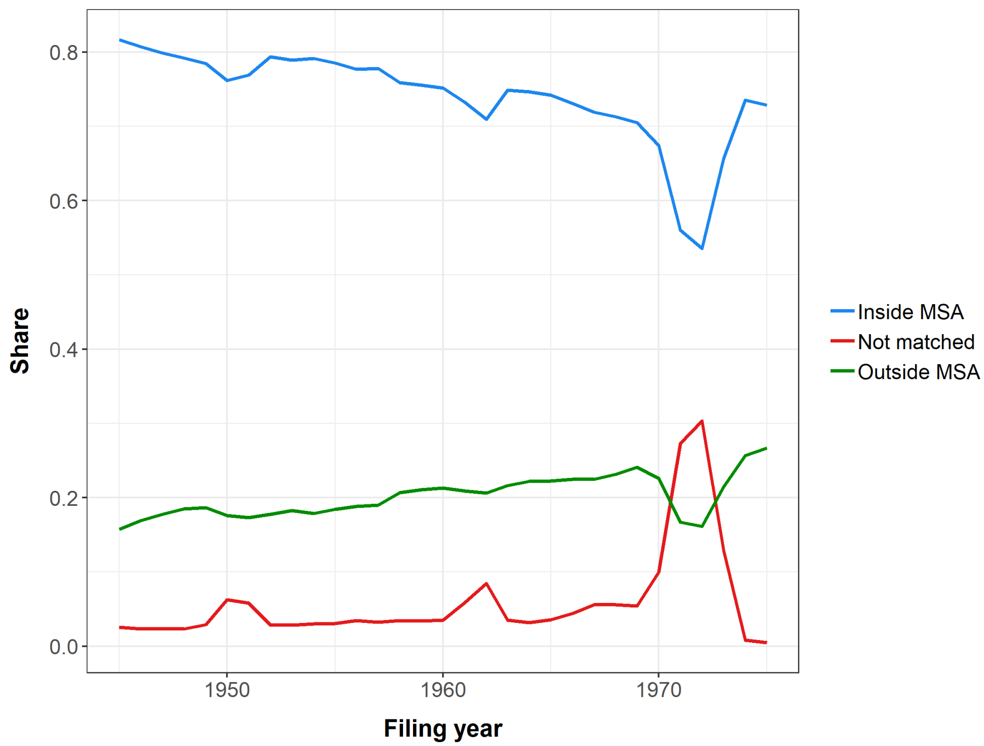


Fig. 1.33. Share patents in Metropolitan Statistical Areas

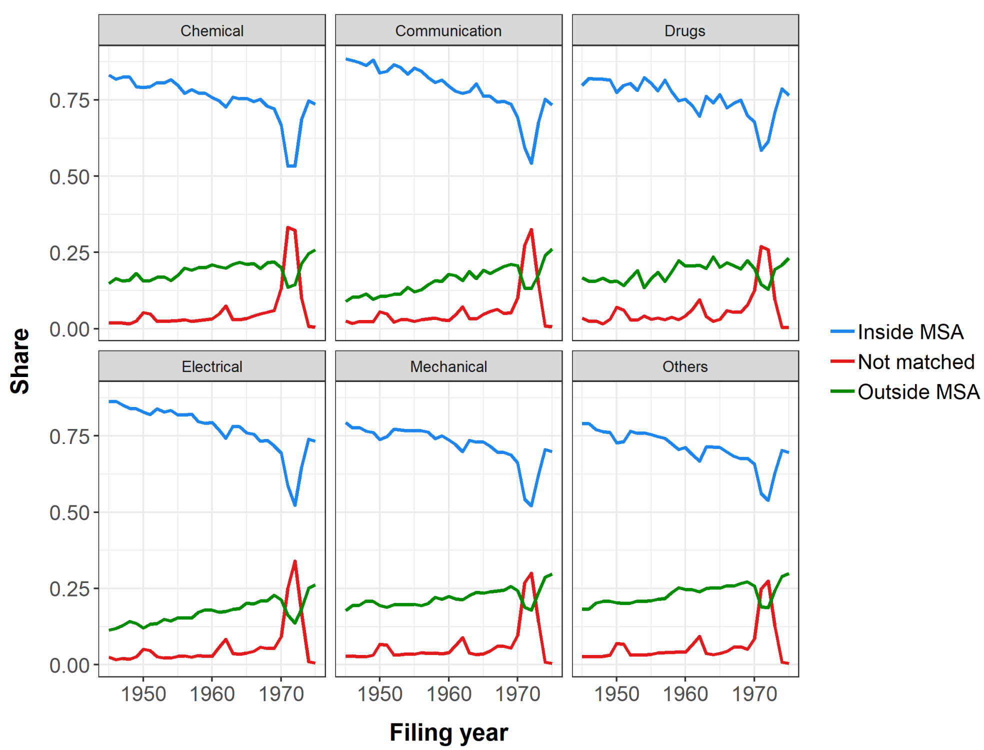


Fig. 1.34. Share patents in Metropolitan Statistical Areas

B.2. Input-Output of patents

In the same spirit as how Input-Output tables of industries are constructed, we can use citations as a reflection of sourced (input) knowledge. In this case, we interpret the cited patent as being a source of knowledge, and the citing patent as being a destination. In Figure 1.35 we aggregate citations by citing-cited technology category in the years 1949-1953. Rows represent the source technology and columns the destination technology. Columns should sum to 1 (round errors may exist). We highlight in bold those IO coefficients that are higher than 0.1. We observe that the diagonal has coefficients greater than 0.5, implying that technologies rely on themselves to create new knowledge. At the same time, we observe the importance of Electrical to create Communication technologies, and the small relevance of Drugs for every other technology.

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**0.74**	0.01	**0.13**	0.03	0.05	0.05
Communication	0	**0.6**	0	0.07	0.01	0.01
Drugs	0.01	0	**0.6**	0	0	0.01
Electrical	0.03	**0.28**	0.03	**0.7**	0.05	0.04
Mechanical	**0.11**	0.07	0.07	0.1	**0.72**	**0.15**
Others	**0.11**	0.05	**0.16**	0.09	**0.16**	**0.75**
Total	1	1	1	1	1	1

Fig. 1.35. Input-Output of technologies 1949-1953

B.3. General Electric research establishments

Using the patent owner identifier we can display the geographical distribution of research establishments for a selected firm. Figure 1.36 shows the research establishments of General Electric in the period 1945-1953. We say that a firm F had a research establishment in location i in time period t if firm F filed at least one patent in time period t with inventors located in location i. The headquarters location q of firm F is defined as the location in which the firm filed the largest amount of patents in the period 1945-1953. General Electric had research establishments in 62 MSAs in the period 1945-1953, and the MSA with the largest amount of patents was Schenectady, New York. Figure 1.37 shows the location of patents cited by patents filed by General Electric with inventors in Fort Wayne, Indiana, in the period 1949-1953. Figure 1.38 shows the research establishments of General Electric during periods 1949-1953 and 1964-1968. General Electric had research establishments in 51 MSAs in 1949-1953 and in 76 MSAs in 1964-1968. 42 out of them appear in both time periods.

Table 1.11: Connected MSAs

MSA fips	MSA name	<=3 periods	4 periods	MSA fips	MSA name	<=3 periods	4 periods
80	Akron, OH SMA	X	X	4680	Macon, GA SMA	X	X
160	Albany-Schenectady-Troy, NY SMA	X	X	4720	Madison, WI SMA	X	X
200	Albuquerque, NM SMA	X	X	4760	Manchester, NH SMA		
240	Allentown-Bethlehem-Easton, PA-NJ SMA	X	X	4920	Memphis, TN SMA	X	X
280	Altoona, PA SMA			5000	Miami, FL SMA	X	X
320	Amarillo, TX SMA	X	X	5080	Milwaukee, WI SMA	X	X
480	Asheville, NC SMA	X		5120	Minneapolis-St. Paul, MN SMA	X	X
520	Atlanta, GA SMA	X	X	5160	Mobile, AL SMA	X	X
560	Atlantic City, NJ SMA	X		5240	Montgomery, AL SMA	X	X
600	Augusta, GA-SC SMA	X	X	5280	Muncie, IN SMA		
640	Austin, TX SMA	X	X	5360	Nashville, TN SMA	X	X
720	Baltimore, MD SMA	X	X	5400	New Bedford, MA SMA		
760	Baton Rouge, LA SMA	X		5440	New Britain-Bristol, CT SMA		
800	Bay City, MI SMA	X		5480	New Haven, CT SMA	X	X
840	Beaumont-Port Arthur, TX SMA	X		5560	New Orleans, LA SMA	X	X
960	Binghamton, NY SMA	X		5600	New York-Northeastern NJ, NY-NJ SMA	X	X
1000	Birmingham, AL SMA	X	X	5720	Norfolk-Portsmouth, VA SMA	X	
1120	Boston, MA SMA	X	X	5840	Ogden, UT SMA	X	
1160	Bridgeport, CT SMA	X	X	5880	Oklahoma City, OK SMA	X	X
1200	Brockton, MA SMA			5920	Omaha, NE-IA SMA	X	X
1280	Buffalo, NY SMA	X	X	5960	Orlando, FL SMA	X	X
1320	Canton, OH SMA	X	X	6120	Peoria, IL SMA	X	
1360	Cedar Rapids, IA SMA	X	X	6160	Philadelphia, PA-NJ SMA	X	X
1440	Charleston, SC SMA	X	X	6200	Phoenix, AZ SMA	X	X
1480	Charleston, WV SMA	X	X	6280	Pittsburgh, PA SMA	X	X
1520	Charlotte, NC SMA	X	X	6320	Pittsfield, MA SMA		
1560	Chattanooga, TN-GA SMA	X	X	6400	Portland, ME SMA		
1600	Chicago, IL-IN SMA	X	X	6440	Portland, OR-WA SMA	X	X
1640	Cincinnati, OH-KY SMA	X	X	6480	Providence, RI SMA	X	X
1680	Cleveland, OH SMA	X	X	6560	Pueblo, CO SMA	X	
1760	Columbia, SC SMA	X	X	6600	Racine, WI SMA	X	X
1800	Columbus, GA-AL SMA	X	X	6640	Raleigh, NC SMA	X	X
1840	Columbus, OH SMA	X	X	6680	Reading, PA SMA	X	X
1880	Corpus Christi, TX SMA	X	X	6760	Richmond, VA SMA	X	X
1920	Dallas, TX SMA	X	X	6800	Roanoke, VA SMA	X	X
1960	Davenport-Rock Island-Moline, IA-IL SMA	X	X	6840	Rochester, NY SMA	X	X
2000	Dayton, OH SMA	X	X	6880	Rockford, IL SMA		
2040	Decatur, IL SMA			6920	Sacramento, CA SMA	X	X
2080	Denver, CO SMA	X	X	6960	Saginaw, MI SMA	X	
2120	Des Moines, IA SMA	X	X	7000	St. Joseph, MO SMA	X	
2160	Detroit, MI SMA	X	X	7040	St. Louis, MO-IL SMA	X	X
2240	Duluth-Superior, MN-WI SMA	X		7160	Salt Lake City, UT SMA	X	X
2280	Durham, NC SMA	X	X	7200	San Angelo, TX SMA		
2320	El Paso, TX SMA	X	X	7240	San Antonio, TX SMA	X	X
2360	Erie, PA SMA	X		7280	San Bernardino, CA SMA		
2440	Evansville, IN SMA	X	X	7320	San Diego, CA SMA	X	X
2480	Fall River, MA-RI SMA	X	X	7360	San Francisco-Oakland, CA SMA	X	X
2640	Flint, MI SMA	X		7400	San Jose, CA SMA		
2760	Fort Wayne, IN SMA	X	X	7520	Savannah, GA SMA	X	
2800	Fort Worth, TX SMA	X	X	7560	Scranton, PA SMA	X	X
2840	Fresno, CA SMA	X	X	7600	Seattle, WA SMA	X	X
2880	Gadsden, AL SMA			7680	Shreveport, LA SMA	X	
2920	Galveston, TX SMA	X	X	7720	Sioux City, IA SMA	X	
3000	Grand Rapids, MI SMA	X		7760	Sioux Falls, SD SMA	X	
3080	Green Bay, WI SMA			7800	South Bend, IN SMA	X	X
3120	Greensboro-High Point, NC SMA	X	X	7840	Spokane, WA SMA	X	X
3160	Greenville, SC SMA	X	X	7880	Springfield, IL SMA	X	
3200	Hamilton-Middletown, OH SMA			7920	Springfield, MO SMA	X	
3240	Harrisburg, PA SMA	X	X	7960	Springfield, OH SMA		
3280	Hartford, CT SMA	X	X	8000	Springfield-Holyoke, MA-CT SMA	X	X
3360	Houston, TX SMA	X	X	8040	Stamford-Norwalk, CT SMA	X	
3400	Huntington-Ashland, WV-KY-OH SMA	X		8120	Stockton, CA SMA	X	X
3480	Indianapolis, IN SMA	X	X	8160	Syracuse, NY SMA	X	X
3520	Jackson, MI SMA	X		8200	Tacoma, WA SMA		
3560	Jackson, MS SMA			8280	Tampa-St. Petersburg, FL SMA	X	X
3600	Jacksonville, FL SMA	X	X	8320	Terre Haute, IN SMA	X	X
3680	Johnstown, PA SMA			8400	Toledo, OH-MI SMA	X	X
3720	Kalamazoo, MI SMA	X		8440	Topeka, KS SMA	X	
3760	Kansas City, MO-KS SMA	X	X	8480	Trenton, NJ SMA		
3800	Kenosha, WI SMA			8560	Tulsa, OK SMA	X	X
3840	Knoxville, TN SMA	X	X	8680	Utica-Rome, NY SMA		
4000	Lancaster, PA SMA	X	X	8800	Waco, TX SMA	X	
4040	Lansing, MI SMA	X		8840	Washington, DC-MD-VA SMA	X	X
4080	Laredo, TX SMA	X		8880	Waterbury, CT SMA		
4160	Lawrence, MA SMA			8920	Waterloo, IA SMA	X	
4280	Lexington, KY SMA	X	X	9000	Wheeling-Steubenville, WV-OH SMA	X	
4320	Lima, OH SMA			9040	Wichita, KS SMA	X	X
4360	Lincoln, NE SMA	X	X	9080	Wichita Falls, TX SMA	X	X

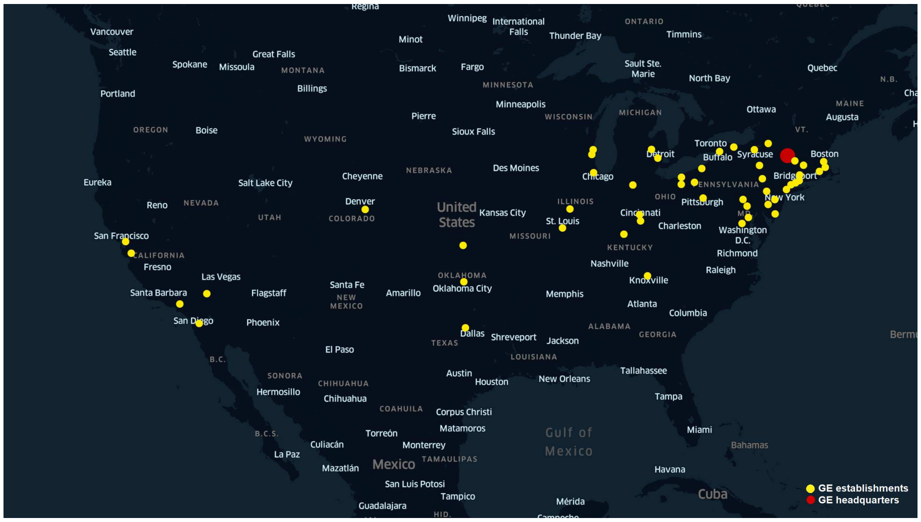


Fig. 1.36. Research establishments of General Electric 1949-1953

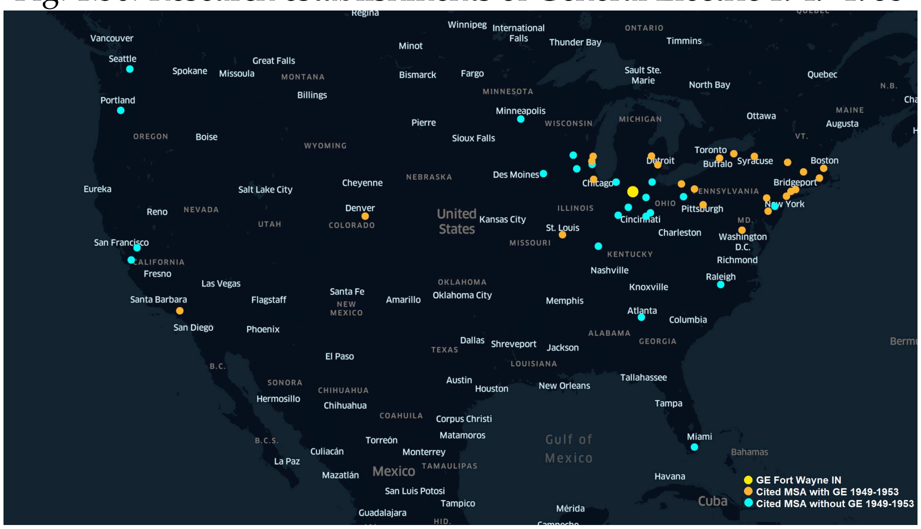


Fig. 1.37. Citations General Electric at Fort Wayne IN 1949-1953

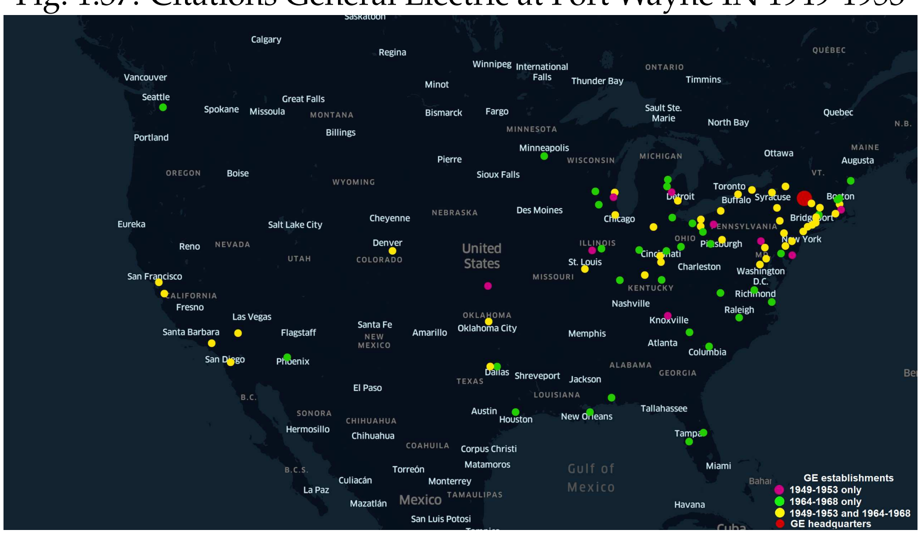


Fig. 1.38. Change location research establishments of General Electric between 1949-1953 and 1964-1968

B.4. *Descriptive statistics*

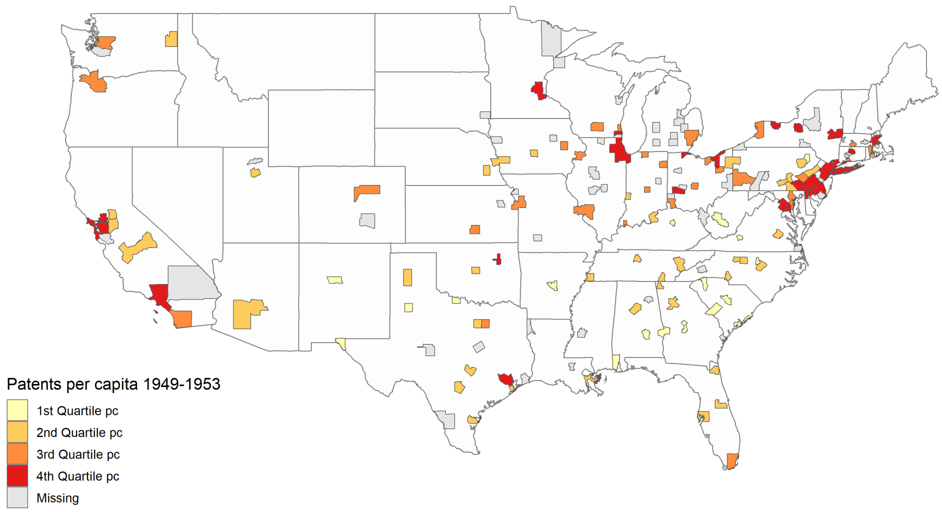


Fig. 1.39. Patents per capita in 1951
Quantiles of patents per capita are computed in each technology and then averaged across technologies. Population is from 1950 Census.

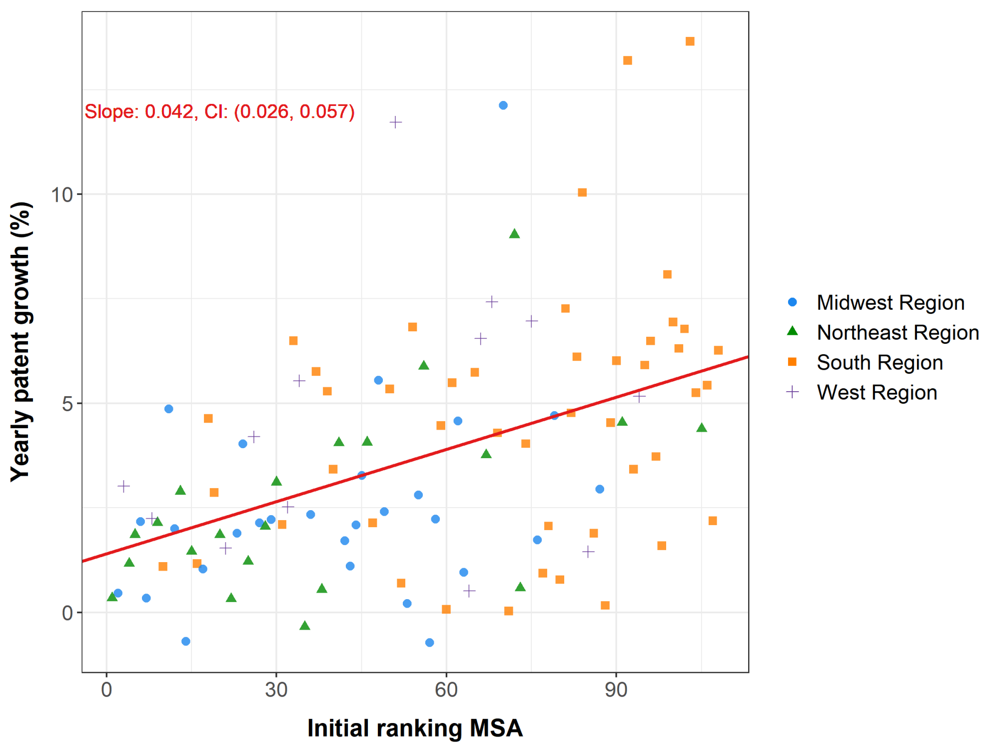


Fig. 1.40. Patent growth by initial innovativeness ranking of MSA

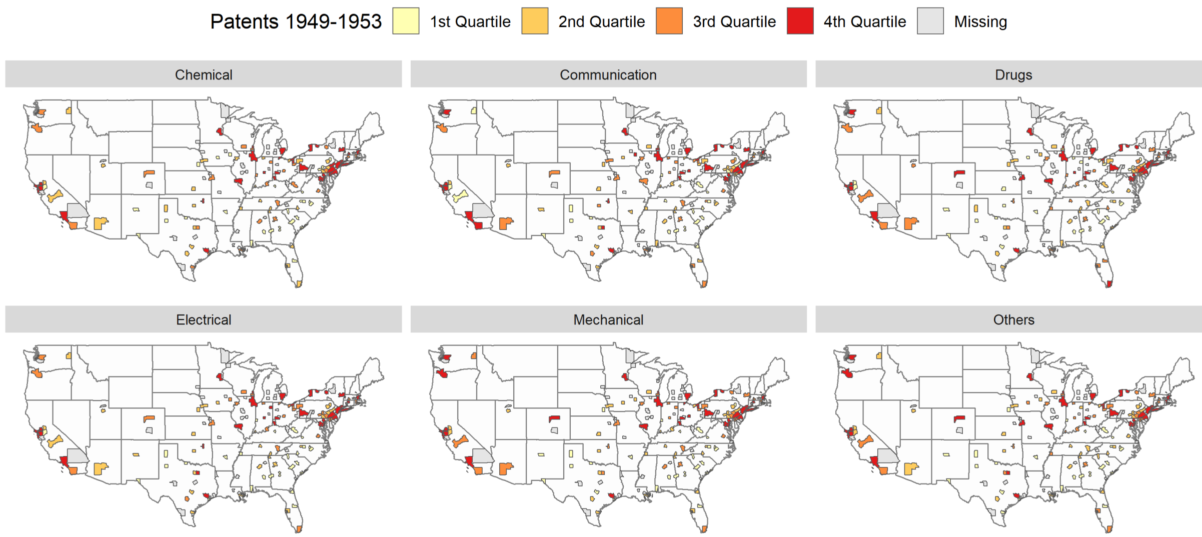


Fig. 1.42. Geography of patenting 1951

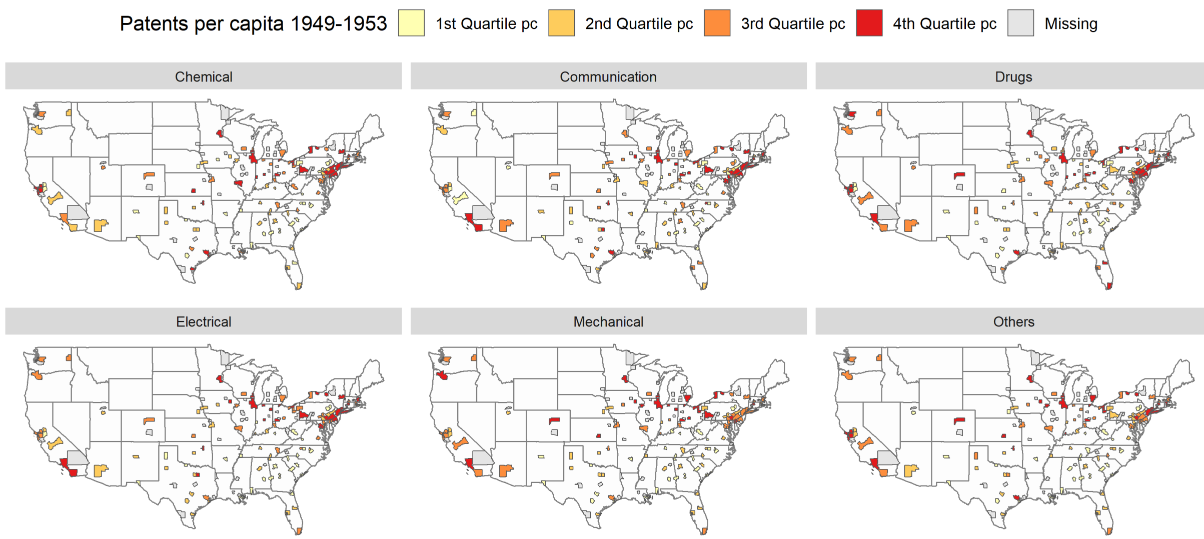


Fig. 1.43. Patents per capita in 1951

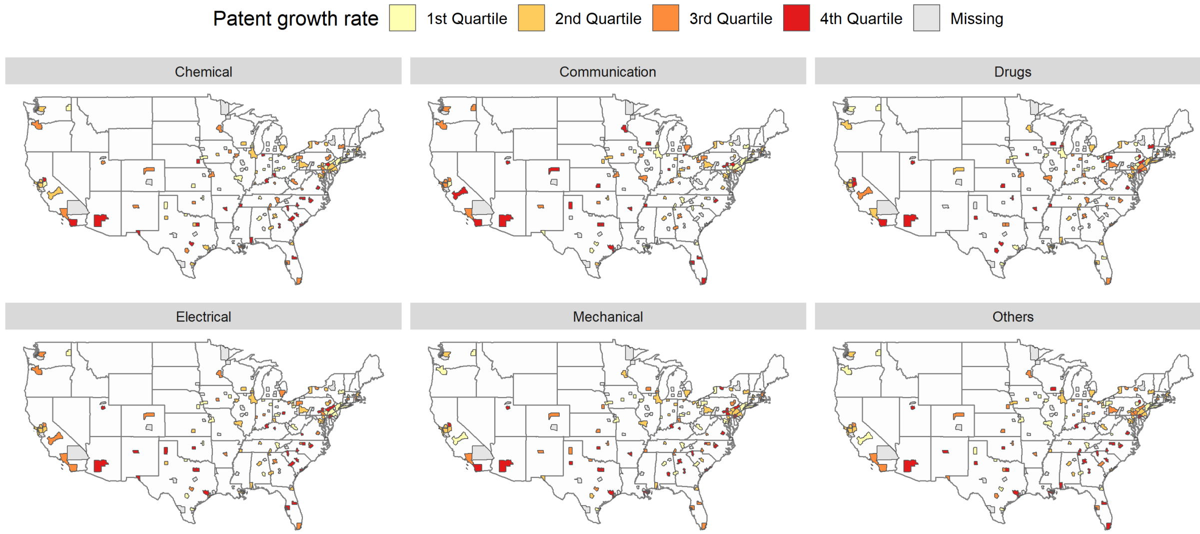


Fig. 1.44. Patent growth rate 1951-1966

B.4.1. Descriptive statistics by technology

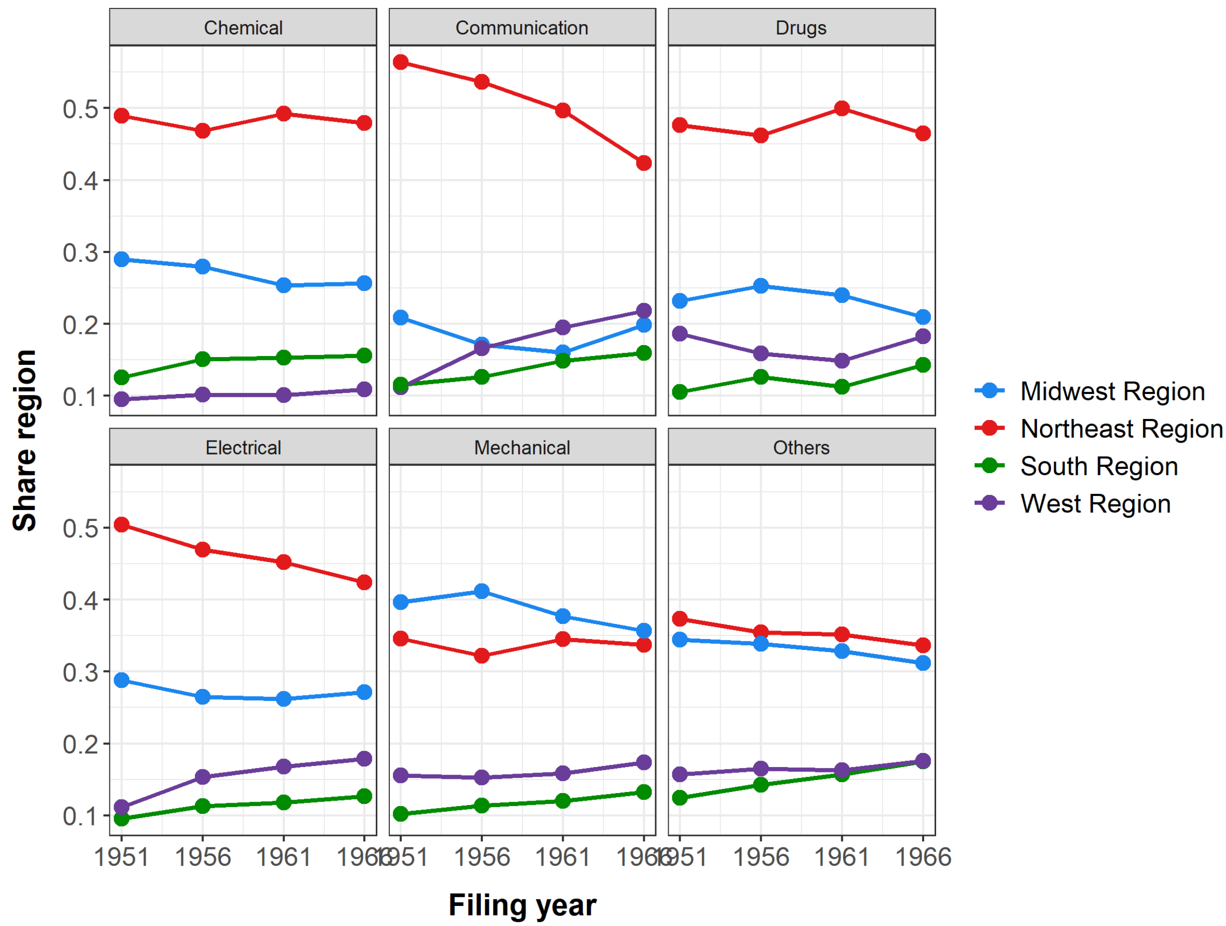


Fig. 1.41. Share of patents by region

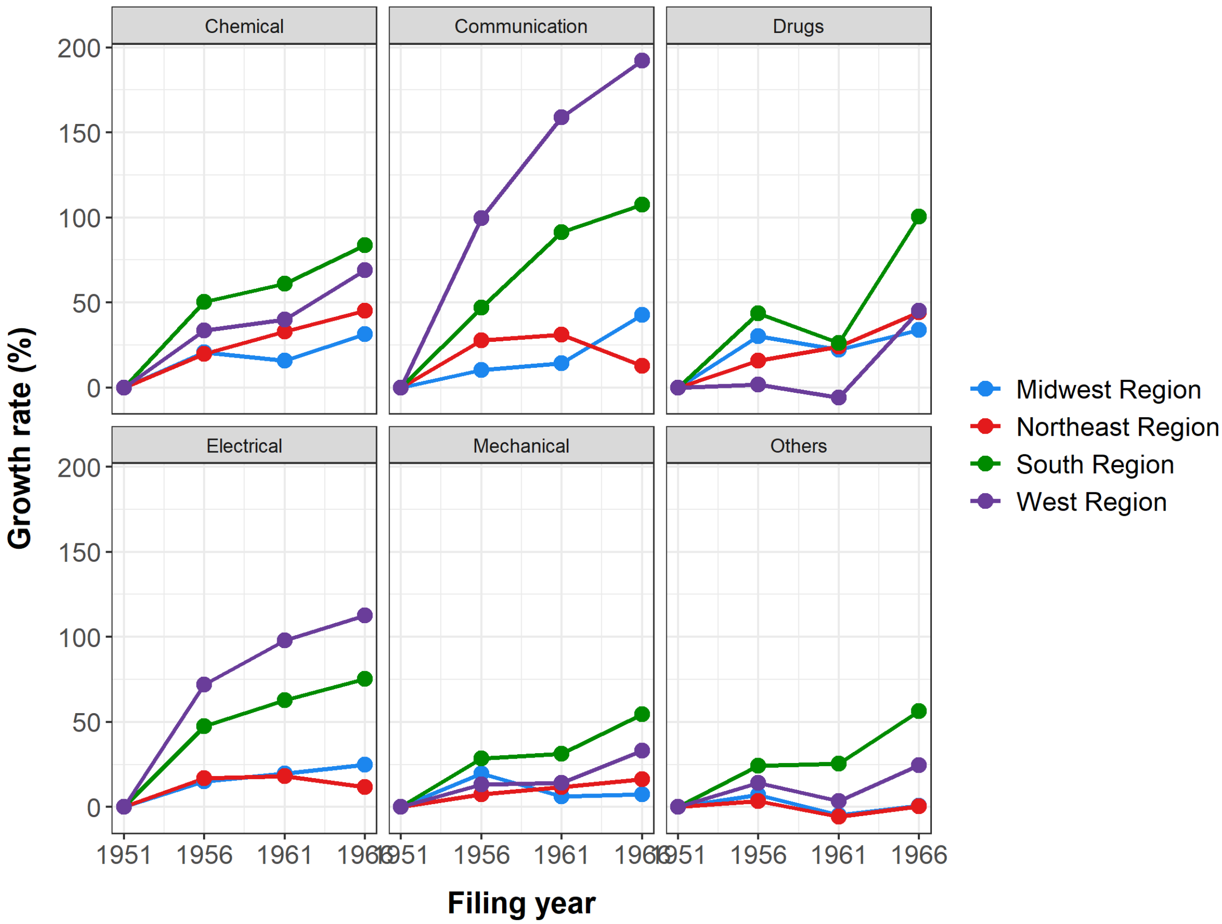


Fig. 1.45. Patent growth rate by region

Technology	Year \ N. estab.	Number of firms: 1	2 to 5	6 to 10	11 to 20	+20	Share of patents: 1	2 to 5	6 to 10	11 to 20	+20
Chemical	1951	6,773	892	72	34	8	0.39	0.25	0.13	0.18	0.06
	1956	7,196	953	108	60	12	0.38	0.22	0.11	0.19	0.10
	1961	6,728	1067	125	80	18	0.32	0.21	0.15	0.15	0.17
	1966	7,092	1120	125	89	30	0.29	0.17	0.13	0.19	0.22
Communication	1951	1,956	270	43	24	8	0.42	0.18	0.18	0.04	0.18
	1956	2,292	337	56	43	11	0.36	0.19	0.15	0.18	0.12
	1961	2,413	441	62	63	15	0.34	0.16	0.09	0.20	0.20
	1966	2,320	414	75	66	29	0.32	0.14	0.08	0.17	0.29
Drugs	1,951	1675	163	20	21	5	0.76	0.19	0.02	0.04	0.00
	1956	1,706	198	40	35	9	0.66	0.18	0.07	0.06	0.02
	1961	1,705	247	57	45	16	0.62	0.19	0.09	0.07	0.04
	1966	2,115	251	49	53	24	0.62	0.13	0.08	0.11	0.06
Electrical	1,951	7394	789	73	33	8	0.47	0.20	0.08	0.08	0.18
	1956	8,182	962	97	59	12	0.44	0.19	0.10	0.12	0.15
	1961	8,077	1,092	123	80	18	0.40	0.19	0.07	0.17	0.17
	1966	7,885	1,006	126	87	30	0.37	0.16	0.09	0.16	0.23
Mechanical	1951	18,509	1,348	75	34	8	0.64	0.20	0.06	0.05	0.04
	1956	18,735	1,498	109	60	12	0.59	0.20	0.06	0.08	0.06
	1961	16,873	1,703	130	80	18	0.54	0.21	0.07	0.10	0.08
	1966	17,856	1,669	132	89	30	0.52	0.17	0.09	0.11	0.11
Others	1951	24,994	1,343	75	34	8	0.76	0.15	0.04	0.03	0.02
	1956	24,650	1,527	110	60	12	0.71	0.16	0.05	0.05	0.03
	1961	20,914	1,683	131	80	18	0.65	0.16	0.06	0.08	0.05
	1966	22,982	1,625	132	89	30	0.63	0.15	0.05	0.08	0.08

Table 1.15: Number of firms and share of patents by firm's geographic coverage

Geographic coverage is computed as the amount of Metropolitan Statistical Areas (MSAs) in which the firm has inventors applying for patents (*research establishments*) in a certain year. Research establishments are defined irrespective of the technology in which technology it patents. Hence a firm applying for patents in technology h in one establishment and technology k in another establishment is defined as having two establishments, and counts as a two-establishment firm both in technology h and technology k. Bins of geographic coverage are 1 MSA, 2 to 5 MSAs, 6 to 10 MSAs, 11 to 20 MSAs, more than 20 MSAs. The maximum possible is 108 MSAs.

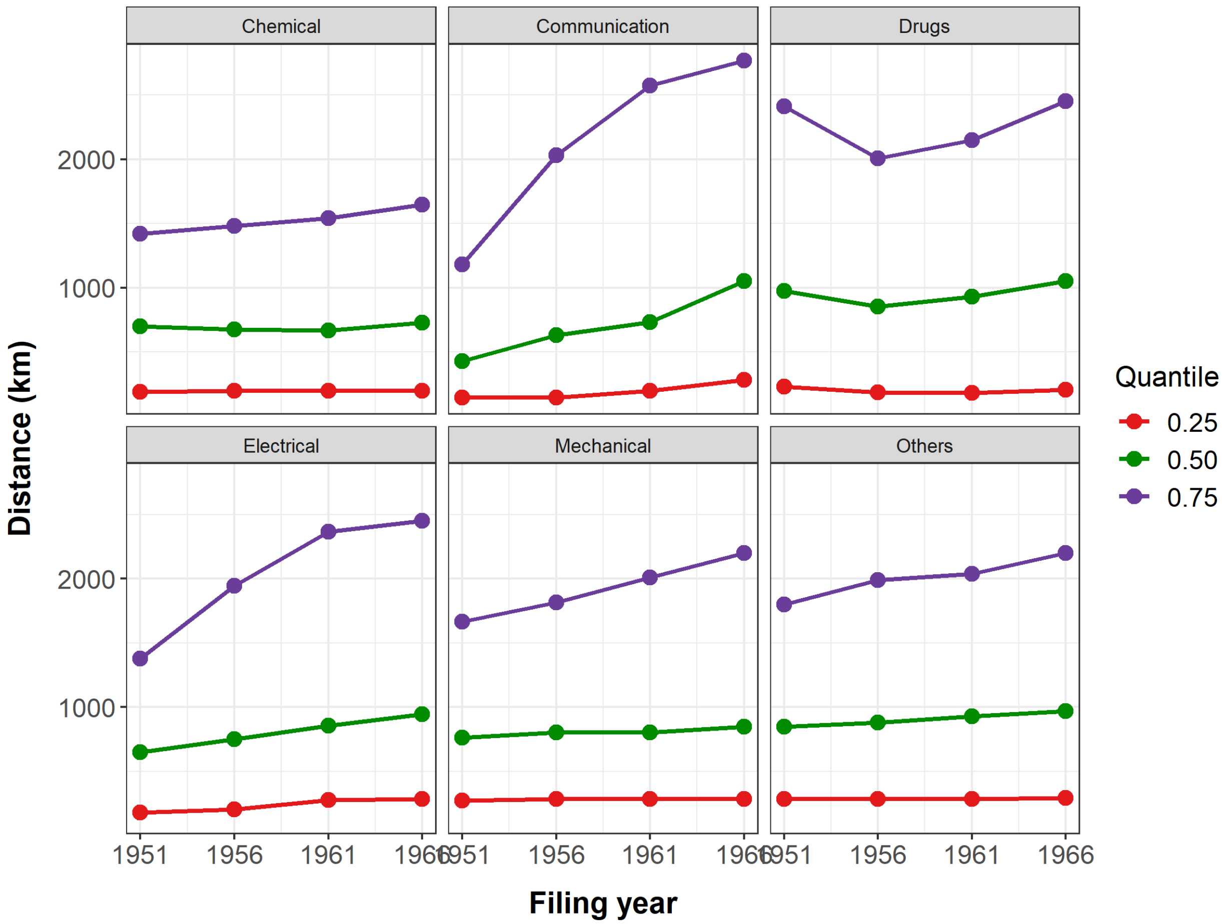


Fig. 1.46. Quantiles of citation distance

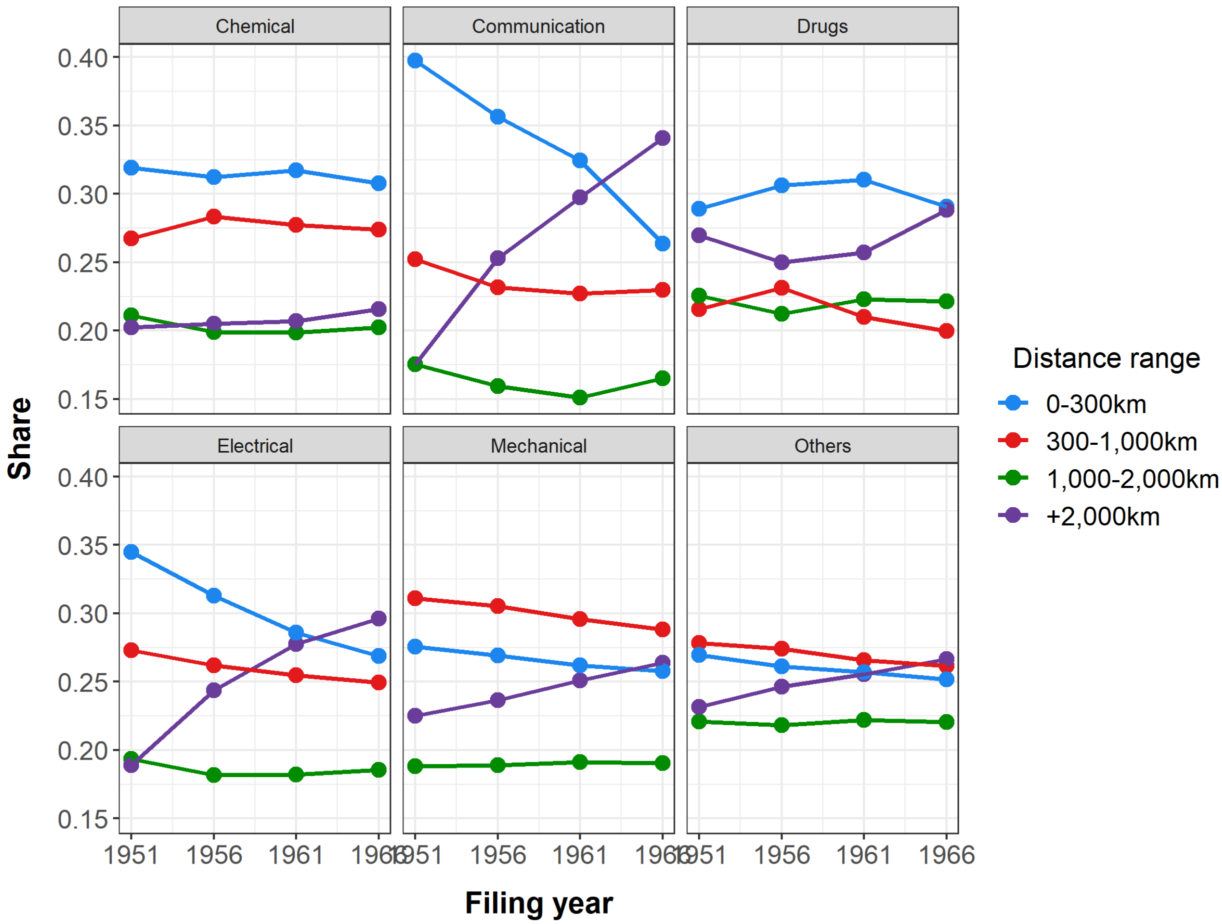


Fig. 1.47. Share of citations by distance

C. US Census Regions

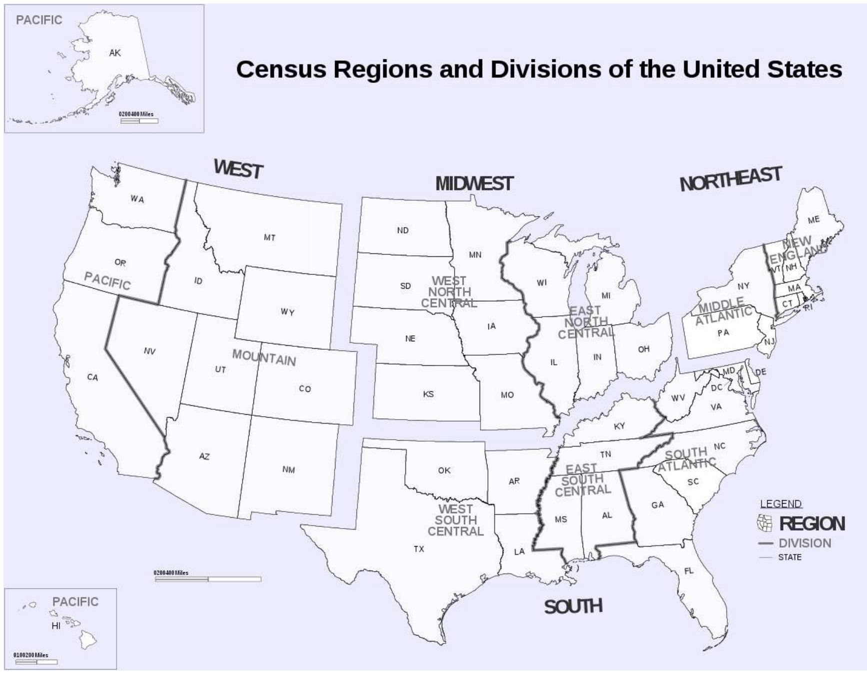


Fig. 1.48. US Census Regions
Source: US Census Bureau

D. Bias Correction and IV estimation

D.1. Split-panel jackknife bias correction

Weidner and Zylkin (2021) show that PPML estimation of gravity equations with three-way fixed effects (origin-time, destination-time, origin-destination) is consistent but asymptotically biased. In their words: *"the asymptotic distribution of the estimates is not centered at the truth as $N \to \infty$"* (page 2). The asymptotic bias concerns both point estimates and standard errors. In order to correct the bias we apply their suggested split-panel jackknife bias correction of section 3.4.1 to both point estimates and bootstrap standard errors. The idea of the jackknife bias correction is to estimate the model in many subsamples and then subtract the average coefficients of the subsamples from (twice) the original coefficient.

As suggested in Weidner and Zylkin (2021) when using real world data (as opposite to simulated data), we estimate the bias correction repeatedly. We modify equation (14) in Weidner and Zylkin (2021) to define the bias corrected coefficient as:

$$\tilde{\beta}_N^J := 2 \times \hat{\beta} - \frac{1}{Z} \sum_z \sum_p \frac{\hat{\beta}_{(p,z)}}{4} \tag{1.12}$$

where p is a random subsample of size 1/4th of the original sample, and Z is the amount of times to subsample.

The *procedure to estimate bias corrected point estimate* $\tilde{\beta}^J_N$ is as follows:

1. Estimate $\hat{\beta}$: the not-bias-corrected estimate of equation ((1.3))
2. Randomly allocate all citing establishment-technology Fih into two equally sized groups (groups are time-invariant). Call them citing groups a and b.
3. Randomly allocate all cited establishment-technology Gjk into two equally sized groups (groups are time-invariant). Call them cited groups a and b.
4. Create four p subsamples of the original data: (a,a), (a,b), (b,a), (b,b). Subsamples keep the same granularity as the original data $FiGjhkt$.
5. Estimate equation ((1.3)) (gravity equation of the main text) in each of the subsamples from the previous step to obtain $\hat{\beta}_{(p,z)}$.[101] Store the four estimated coefficients.
6. Repeat Z times steps 2 to 5.
7. Compute equation (1.12)

To compute bias-corrected bootstrap standard errors we need to bias-correct the point estimate $\tilde{\beta}^J_m$ of each bootstrap iteration m. The *procedure to estimate bias corrected standard errors* is as follows:

1. Sample establishment-technology-pairs $FiGjhk$ with replacement such that we obtain a re-sampled data of the same size as the original data (hence, some $FiGjhk$ will be repeated in the re-sampled data). Sampled $FiGjhk$ are kept for all time periods in order to keep the source of identification of β: across time variation within a establishment pair. Label this new dataset $data_m$.
2. Using $data_m$, estimate equation ((1.3)) to obtain $\hat{\beta_m}$ (this is a point estimate of the specific $data_m$)
3. Using $data_m$, repeat Z_M times steps 2 to 5 of the *procedure to estimate bias corrected point estimate*. This step provides $Z_M \times 4$ point estimates $\hat{\beta}_{(p,m,z_M)}$
4. Compute the bias corrected point estimate of bootstrap m $\tilde{\beta}^J_m = 2 \times \hat{\beta}_m - \frac{1}{Z_M} \sum_{z_M} \sum_p \frac{\hat{\beta}_{(p,m,z_M)}}{4}$.
5. Store the bias corrected point estimate of bootstrap m
6. Repeat steps 1 to 5 M times to obtain M bias corrected bootstrap point estimates $\tilde{\beta}^J_m$
7. Compute the variance-covariance matrix of bias corrected bootstrap coefficients $\tilde{\beta}^J_m$ and use it to compute standard errors of $\tilde{\beta}^J_N$

The bias correction of point estimates and bias correction of bootstrap standard errors implies estimating $Z \times 4 + Z_M \times M \times 4$ models. This is a computationally demanding task. To estimate columns (1) and (2) of Table 1.2 we set $Z = 100$, $Z_M = 5$ and $M = 200$, adding up to $1,100$ models to estimate for each column.

[101] Given that we require to identify the fixed effects, the *effective subsample* in all four p estimations does not have the same amount of observations. However, in our estimations the *effective subsample* size across p subsamples does not differ by more than 5%.

As recommended in Hansen (2021), in the Table 1.16 we repeat Table 1.2 but reporting 0.025 and 0.975 quantile values of bootstrap estimates (bias corrected for columns (1) and (2)) instead of standard errors:

	PPML		IV PPML	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$cit_{FiGjhkt}$	
	(1)	(2)	(3)	(4)
log(travel time)	-0.083 $(-0.129; -0.056)$		-0.152 $(-0.210; -0.097)$	
log(travel time) $\times$ 0-300km		0.019 $(-0.054; 0.082)$		-0.076 $(-0.542; 0.384)$
log(travel time) $\times$ 300-1,000km		-0.089 $(-0.141; -0.052)$		-0.134 $(-0.246; -0.066)$
log(travel time) $\times$ 1,000-2,000km		-0.094 $(-0.156; -0.022)$		-0.112 $(-0.192; -0.022)$
log(travel time) $\times$ +2,000km		-0.169 $(-0.277; -0.105)$		-0.203 $(-0.311; -0.136)$
N obs. effective	$4,703,010$	$4,703,010$	$4,703,010$	$4,703,010$
R2	0.88	0.88	0.88	0.88

Table 1.16: Elasticity of citations to travel time

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (2) includes the interaction of travel time$_{ijt}$ with a dummy for distance bin between the citing establishment Fi and the cited establishment Gj. Column (3) and (4) show the result of two step instrumental variables estimation, where $log(\text{travel time}_{ijt})$ is instrumented with $log(\text{travel time}_{ijt}^{\text{fix routes}})$, the travel time that would have taken place if routes were fixed to the ones observed in 1951 and in each year routes were operated with the average airplane of the year. 0.025 and 0.975 quantile bootstrap estimates are presented in parentheses. The coefficients and bootstrap estimates in columns (1) and (2) are jackknife bias-corrected. R2 is computed as the squared correlation between observed and fitted values.

D.2. *Instrumental variables PPML*

To implement the instrumental variables of Poisson estimation we follow the control function approach described in Wooldridge (2014). We explain the procedure using the estimation of the elasticity of citations to travel time. The procedure is similar for the elasticity of (new) patents to knowledge access. We proceed in two steps estimating the following two equations:

$$\begin{aligned} log(\text{travel time})_{FiGjhkt} &= \lambda_2\, log(\text{travel time}_{FiGjhkt}^{\text{fix routes}}) \\ &\quad + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + u_{FiGjhkt} \end{aligned} \tag{1.13}$$

$$citations_{FiGjhkt} = \exp\left[\beta \, log(\text{travel time}_{ijt}) + \lambda \, \hat{u}_{FiGjhkt} + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times v_{FiGjhkt} \tag{1.14}$$

In a first step we estimate equation ((1.13)) and obtain estimated residuals $\hat{u}_{FiGjhkt}$. In a second step we use the estimated residuals as a regressor in equation ((1.14)) which *controls* for the endogenous component of travel time.

To perform inference we bootstrap standard errors in the following way:

1. Sample establishment-technology-pairs $FiGjhk$ with replacement such that we obtain a re-sampled data of the same size as the original data (hence, some $FiGjhk$ will be repeated in the re-sampled data). Sampled $FiGjhk$ are kept for all time periods in order to keep the source of identification of β: across time variation within a establishment pair. Label this new dataset $data_m$
2. Using $data_m$, estimate equations ((1.13)) and ((1.14)) to obtain the bootstrap estimate $\hat{\beta}_m$. Store $\hat{\beta}_m$.
3. Repeat M times steps 1 and 2.
4. Compute the variance-covariance matrix of $\hat{\beta}_m$ and use it to compute standard errors of $\hat{\beta}$

For columns (3) and (4) of Table 1.2, and columns (3) and (4) of Table 1.4 we set $M = 200$.

E. Additional results

E.1. Diffusion of knowledge

E.1.1. Heterogeneous effects

First, we perform an intensive margin/extensive margin decomposition of the effect of travel time on citations. We find that the effect is coming from both margins. In the instrumental variables approach, the intensive margin is only statistically different from zero for distance greater than 2,000km, while for the extensive margin it is for distance greater than 300km. Results for the baseline analysis are shown in Table 1.17 and for the IV estimation in Table 1.18.

Second, we investigate if the elasticity varies by the degree of concentration of patents across establishments in the citing technology or cited technology, we find no statistically significant heterogeneous effect. Results are shown in columns (1) and (2) of Table 1.20.

Third, we check if the elasticity varies by the median forward and backward citation lags of the cited and citing technologies. We find that the elasticity of citations to travel time is *more negative* both for technologies that accumulate citations during a longer time period and for technologies that cite older patents. To be able to precisely show if it is

newer or *older* technologies that diffuse better as consequence of the jet requires an analysis with the citation level forward and backward lag, and not using the median lag in the technology. Nonetheless, the results seem to suggest that jets improved the diffusion of *older* technologies. Results are shown in columns (3) and (4) of Table 1.20.

Fourth, we extend the sample of patents to include patents with a patent owner identified as a government organization or university. Column (5) of Table 1.20 opens the elasticity of citations to travel time by whether the citing patent belongs to a government organization of university. Column (6) includes a dummy for whether the cited patent belongs to a government organization or university. We do not observe a particular change in the pattern of the elasticity of citations to travel time.

Sixth, we extend the sample to include self citations (citations in which the citing and cited patents belong to the same patent owner F). Column (7) of Table 1.20 shows that the elasticity is not statistically different for self citations.

Seventh, we check if the elasticity varies with the level of innovativeness of the citing firm. It may be the case that those firms that actually have the -time and monetary- budget to take a plane are only the most innovative ones. We rank firms F in technology h according to the amount of patents filed by F in technology h at the initial time period 1949-1953. We define quantile 0.00 as all those firms that did not file patents in 1949-1953, while quantile 0.01 is assigned to those that filed patents but not as many as to be in the quantile 0.25 or higher. Results are shown in Table 1.19. We do not find a particular pattern related to the initial innovativeness.

Eighth, we check if the elasticity varies with the citing technology, cited technology and citing-cited technology pair. Results are shown in Table 1.21 and Table 1.24. We find that the elasticity is negative and significant mainly when the citing and cited technology are the same. In Appendix B we show that most citations happen within a technology, so most identification power would be when citing and cited technologies are the same.

	PPML		log-log		linear probability	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$log(cit_{FiGjhkt})$		$cit_{FiGjhkt} > 0$	
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.083^{***} (0.019)		-0.071 (0.098)		-0.013*** (0.003)	
log(travel time):0-300km		0.019 (0.036)		0.318^{**} (0.152)		-0.0045 (0.005)
log(travel time):300-1000km		-0.089^{***} (0.023)		-0.265^{*} (0.145)		-0.008^{***} (0.003)
log(travel time):1000-2000km		-0.094^{***} (0.032)		-0.231 (0.209)		-0.013^{***} (0.003)
log(travel time):+2000km		-0.169^{***} (0.039)		-0.424^{**} (0.192)		-0.024^{***} (0.005)
N obs. effective	4,703,010	4,703,010	16,412	16,412	10,106,940	10,106,940
R2	0.88	0.88	0.86	0.86	0.70	0.70

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.17: Elasticity of citations to travel time: intensive and extensive margin

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\beta\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (3) shows the result of an OLS estimation of $log(citations_{FiGjhkt}) = \alpha\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with a sample of establishment-technology pairs ($FiGjhk$) that have positive citations in all periods. Column (5) shows the result of an OLS estimation of $\mathbb{1}\{citations_{FiGjhkt} > 0\} = \gamma\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with the same sample as (1). Column (2), (4) and (6) open, respectively, the coefficients β, α, γ by distance between the citing establishment Fi and the cited establishment Gj. Standard errors are presented in parentheses. Columns (1) and (2) present coefficients and bootstrap standard errors jackknife bias corrected. Columns (3) through (6) present standard errors clustered at the non-directional location pair (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	IV PPML		IV log-log		IV linear probability	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$log(cit_{FiGjhkt})$		$cit_{FiGjhkt} > 0$	
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.152^{***} (0.029)		-0.396** (0.175)		-0.027*** (0.004)	
log(travel time):0-300km		-0.076 (0.221)		1.324 (1.680)		-0.028 (0.036)
log(travel time):300-1000km		-0.134^{***} (0.044)		-0.148 (0.378)		-0.022*** (0.007)
log(travel time):1000-2000km		-0.112^{**} (0.047)		-0.314 (0.200)		-0.021*** (0.005)
log(travel time):+2000km		-0.203^{***} (0.043)		-0.388** (0.185)		-0.032*** (0.005)
N obs. effective	4,703,010	4,703,010	16,412	16,412	10,106,940	10,106,940
R2	0.88	0.88	0.86	0.86	0.70	0.70

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.18: Elasticity of citations to travel time: IV estimation intensive and extensive margin

Column (1) shows the result of Instrumental Variables Poisson estimation of $citations_{FiGjhkt} = \exp\left[\beta\, log(\text{travel time}_{ijt}) + \lambda\, \hat{u}_{FiGjhkt} + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. The variable $\hat{u}_{FiGjhkt}$ is constructed as $\hat{u}_{FiGjhkt} = \text{travel time}_{FiGjhkt} - \hat{\lambda}_2\, \text{travel time}^{\text{fix network}}_{FiGjhkt}$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (3) shows the result of an IV-2SLS estimation of $log(citations_{FiGjhkt}) = \alpha\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with a sample of establishment-technology pairs ($FiGjhk$) that have positive citations in all periods. Column (5) shows the result of an IV-2SLS estimation of $\mathbb{1}\{citations_{FiGjhkt} > 0\} = \gamma\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with the same sample as (1). Columns (3) and (5) use travel time$^{\text{fix network}}_{ijt}$ as an instrument for travel time$_{ijt}$. Column (2), (4) and (6) open, respectively, the coefficients β, α, γ by distance between the citing establishment Fi and the cited establishment Gj. Standard errors are presented in parenthesis. In Columns (1) and (2) standard errors are bootstrapped. In Columns (3) to (6) standard errors clustered at the non-directional location pair (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	Concentration citing	Concentration cited	Cited lag forward	Citing lag backward	Citing govnt & uni	Cited govnt & univ	Self citation
Dep. variable: *citations*				$cit_{FiGjhkt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(travel time):0-300km	0.103	0.160	−0.045	0.1907	0.021	0.018	0.002
	(0.121)	(0.114)	(0.472)	(0.538)	(0.038)	(0.038)	(0.039)
log(travel time):300-1000km	−0.105	−0.039	−0.546	−0.145	-0.102^{***}	-0.099^{***}	-0.077^{***}
	(0.084)	(0.095)	(0.364)	(0.366)	(0.027)	(0.027)	(0.029)
log(travel time):1000-2000km	−0.138	−0.117	0.086	0.101	-0.094^{**}	-0.093^{**}	-0.094^{**}
	(0.105)	(0.116)	(0.480)	(0.498)	(0.042)	(0.041)	(0.040)
log(travel time):+2000km	-0.287^{***}	-0.268^{***}	0.720^{**}	0.560	-0.185^{***}	-0.188^{***}	-0.153^{***}
	(0.105)	(0.090)	(0.344)	(0.472)	(0.049)	(0.048)	(0.040)
log(travel time):0-300km × X	−1.180	−2.013	0.028	−0.066	−0.125	0.481	0.038
	(1.843)	(1.712)	(0.185)	(0.211)	(0.367)	(0.543)	(0.252)
log(travel time):300-1000km × X	0.079	−0.880	0.178	0.018	−0.088	-0.609^{*}	0.077
	(1.188)	(1.366)	(0.144)	(0.145)	(0.265)	(0.330)	(0.127)
log(travel time):1000-2000km × X	0.634	0.341	−0.073	−0.078	−0.282	−0.370	0.082
	(1.412)	(1.606)	(0.191)	(0.197)	(0.366)	(0.385)	(0.210)
log(travel time):+2000km × X	1.436	1.157	-0.366^{***}	−0.299	−0.328	0.015	−0.073
	(1.456)	(1.136)	(0.137)	(0.188)	(0.410)	(0.295)	(0.170)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,800,144	4,800,144	4,835,001
R2	0.88	0.88	0.88	0.88	0.88	0.88	0.94

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.20: Elasticity of citations to travel time: Heterogeneity (part 1)

Result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp[\sum_d \beta_d \mathbb{1}\{distance_{ij} \in d\}\, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \mathbb{1}\{distance_{ij} \in d\}\, \mathbb{1}\{X_{FiGjhkt}\}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0-300km]$, $(300km-1000km]$, $(1000km-2000km]$, $(2000km-max]$. The variable X takes different value depending on the column: in column (1) it is the across-MSA Herfindahl index of the citing technology, in column (2) it is the across-MSA Herfindahl index of the cited technology, in column (3) it is median forward citation lag of the cited technology, in column (4) it is median backward citation lag of the citing technology. In column (5) and (6) the sample includes government and university patents, in column (5) X is a dummy that takes value one if the citing patent belongs to a university or government organisation, in column (6) it is a dummy that takes value one if the cited patent belongs to a university or government organisation. In column (7) the sample includes self citations, the variable X is a dummy that takes value one if the citing firm F cited firm G are the same. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	Citing quantile	Cited quantile
Dep. variable: *citations*	$cit_{FiGjhkt}$	
	(1)	(2)
log(travel time) × quantile 0.00	-0.151***	-0.111***
	(0.058)	(0.039)
log(travel time) × quantile 0.01	-0.078	-0.084
	(0.114)	(0.101)
log(travel time) × quantile 0.25	-0.081	-0.159*
	(0.103)	(0.093)
log(travel time) × quantile 0.50	-0.139	-0.063
	(0.091)	(0.083)
log(travel time) × quantile 0.75	-0.262***	-0.033
	(0.079)	(0.068)
log(travel time) × quantile 0.90	-0.029	-0.127**
	(0.066)	(0.057)
log(travel time) × quantile 0.95	-0.001	-0.123***
	(0.037)	(0.038)
log(travel time) × quantile 0.99	-0.130***	-0.066*
	(0.035)	(0.039)
log(travel time) × quantile 0.999	-0.070	-0.070
	(0.045)	(0.045)
N obs. effective	4,703,010	4,703,010
R2	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.19: Elasticity of citations to travel time: Heterogeneity (part 2)

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_q \boldsymbol{\beta}_q \, log(\text{travel time}_{ijt})\, \mathbb{1}\{quantile_{Fh} \in q\} + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time_{ijt} is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. $quantile_{Fh}$ is the quantile of firm F in the distribution of firms within technology h, using patents applied by F in h in the time period 1949-1953. Column (2) repeats the analysis using the quantile of the cited firm G in technology k. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	PPML	
	Citing technology	Cited technology
Dep. variable: *citations*	$cit_{FiGjhkt}$	
	(1)	(2)
log(travel time) × Chemical	-0.066	-0.093^{**}
	(0.045)	(0.045)
log(travel time) × Computers & Communications	-0.100	-0.140^{*}
	(0.079)	(0.077)
log(travel time) × Drugs & Medical	-0.053	-0.005
	(0.162)	(0.181)
log(travel time) × Electrical & Electronic	-0.070	-0.054
	(0.048)	(0.046)
log(travel time) × Mechanical	-0.080^{**}	-0.087^{***}
	(0.031)	(0.032)
log(travel time) × Others	-0.147^{***}	-0.113^{**}
	(0.045)	(0.044)
N obs. effective	4,703,010	4,703,010
R2	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.21: Elasticity of citations to travel time by citing and cited technology
Part 1

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_{tech} \beta_h \mathbb{1}\{tech = h\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G located in j, in technology k. $\mathbb{1}\{tech = h\}$ is a dummy variable that takes value 1 when the citing technology h is equal to technology $tech$. In column (2) the dummy is modified to $\mathbb{1}\{tech = k\}$ such that it takes value 1 when the cited technology k is equal to technology $tech$. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

E.1.2. IV PPML: first and second stage estimation

	First stage OLS	Second stage PPML
Dep. variable:	log(travel time)	$cit_{FiGjhkt}$
	(1)	(2)
log(travel time fix routes)	0.951^{***} (0.039)	
log(travel time)		-0.152^{***} (0.029)
residual		0.094^{***} (0.035)
N obs. effective	10,106,940	4,703,010
R2	0.99	0.88
Within R2	0.38	

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.22: Elasticity of citations to travel time: first and second stage IV PPML

	OLS First stage 0-300km	OLS First stage 300-1,000km	OLS First stage 1,000-2,000km	OLS First stage +2,000km	Second stage PPML
Dep. variable:	log(travel time)				$cit_{FiGjhkt}$
	(1)	(2)	(3)	(4)	(5)
log(travel time fix routes) × 0-300km	0.278^{**} (0.122)	0.073 (0.057)	0.024 (0.026)	0.040^{*} (0.022)	
log(travel time fix routes) × 300-1,000km	-0.103^{***} (0.032)	1.113^{***} (0.041)	-0.013 (0.011)	0.010 (0.011)	
log(travel time fix routes) × 1,000-2,000km	-0.064^{***} (0.024)	-0.052^{***} (0.020)	1.059^{***} (0.044)	0.017^{*} (0.009)	
log(travel time fix routes) × +2,000km	-0.058^{***} (0.022)	-0.046^{***} (0.017)	-0.020^{**} (0.010)	1.097^{***} (0.018)	
log(travel time) × 0-300km					-0.076 (0.221)
log(travel time) × 300-1,000km					-0.134^{***} (0.044)
log(travel time) × 1,000-2,000km					-0.112^{**} (0.047)
log(travel time) × +2,000km					-0.203^{***} (0.043)
residual × 0-300km					0.100 (0.196)
residual × 300-1,000km					0.045 (0.053)
residual × 1,000-2,000km					0.026 (0.069)
residual × +2,000km					0.043 (0.078)
N obs. effective	10,106,940	10,106,940	10,106,940	10,106,940	4,703,010
R2	0.99	0.99	0.99	0.99	0.88
Within R2	0.04	0.46	0.80	0.88	

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.23: Elasticity of citations to travel time: first and second stage IV PPML

Cited \ Citing	Chemical	Computers & Communications	Drugs & Medical	Electrical & Electronic	Mechanical	Others
Chemical	-0.092^{**}	0.219	0.113	-0.299^{***}	-0.025	-0.070
	(0.052)	(0.262)	(0.199)	(0.094)	(0.071)	(0.068)
Computers & Communications	-0.089	-0.306^{***}	-0.657	0.107	0.122	0.095
	(0.259)	(0.095)	(0.976)	(0.090)	(0.149)	(0.169)
Drugs & Medical	0.224	0.567	-0.278	-0.230	-0.334	0.358
	(0.239)	(1.205)	(0.268)	(0.561)	(0.362)	(0.323)
Electrical & Electronic	0.233^{**}	0.171^{*}	-0.224	-0.102^{**}	0.087	-0.063
	(0.093)	(0.096)	(0.634)	(0.056)	(0.070)	(0.079)
Mechanical	-0.060	0.151	-0.152	0.106	-0.129^{***}	-0.032
	(0.076)	(0.145)	(0.402)	(0.082)	(0.035)	(0.056)
Others	0.042	0.173	0.204	0.052	0.019	-0.209^{***}
	(0.074)	(0.169)	(0.274)	(0.072)	(0.053)	(0.054)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.24: Elasticity of citations to travel time by citing and cited technology
Part 2

Column (1) shows the result of one single Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_{\text{tech pair}} \beta_{hk}\, \mathbb{1}\{\text{tech pair} = hk\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G located in j, in technology k. $\mathbb{1}\{\text{tech pair} = hk\}$ is a dummy variable that takes value 1 when the citing technology h is equal to technology $tech$. In column (2) the dummy is modified to $\mathbb{1}\{tech = k\}$ such that it takes value 1 when the cited technology k is equal to technology $tech$. travel time_{ijt} is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values. The amount of observation in the effective sample is 4,703,010.

E.1.3. Robustness

Sample of establishments

During the time period there was entry and exit of research establishments that was not uniform across locations. We may then think that the change in diffusion of knowledge is only consequence of the change in the geographical location of innovation. To test this possibility, in Table 1.25 we estimate the baseline regression (1.3) with different samples. In column (1) we include the baseline results.[102] In column (2) we use only citing establishments Fi that filed patents during the initial time period 1949-1953. In column (3) we further restrict the sample to both citing establishments Fi and cited establishments Gj that

[102] Coefficients are not bias corrected.

filed patents in 1949-1953.[103] We find that the coefficient at more than 2,000km remains comparable to the one in the baseline regression, statistically significant at the 1%.

[103]We require Fi and Gj to have positive amount of patents applied during 1949-1953. However, those establishments need not to have cited each other.

	All	Citing establishment	Citing & Cited establishment
Dep. variable: *citations*		$cit_{FiGjhkt}$	
	(1)	(2)	(3)
log(travel time) × 0-300km	0.021	0.020	0.028
	(0.039)	(0.043)	(0.043)
log(travel time) × 300-1,000km	-0.099^{***}	-0.095^{***}	-0.095^{***}
	(0.027)	(0.029)	(0.030)
log(travel time) × 1,000-2,000km	-0.093^{**}	-0.092^{**}	-0.062
	(0.042)	(0.047)	(0.050)
log(travel time) × +2,000km	-0.185^{***}	-0.155^{***}	-0.179^{***}
	(0.049)	(0.052)	(0.052)
N obs. effective	4,703,010	3,109,285	1,960,851
R2	0.88	0.88	0.89

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.25: Elasticity of citations to travel time: Fix sample of establishments

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \times \mathbb{1}\{distance_{ij} \in d\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km]$, $(300km - 1000km]$, $(1000km - 2000km]$, $(2000km - max]$. Column (2) truncates the sample keeping only citing establishments Fi that where present in the initial time period $1949 - 1953$. Column (3) truncates the sample keeping only citing establishments Fi and cited establishments Gj that where present in the initial time period. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Ticket prices

During the period of analysis ticket prices were set by the Civil Aeronautics Board, so airlines could not set prices of their own tickets. Some airlines included a sample of prices in the last page of their booklet of flight schedules a sample of prices, which we digitized. We have digitized American Airlines 1951, 1961, 1966; TWA 1951 and United Airlines 1956 and 1961.[104]. The sample includes prices for 11,590 directional airport pair years. We document multiple facts about prices.

First, prices were set in the form of an intercept plus a variable increment depending on distance between origin and destination (until 1962-1963). A linear regression with an intercept and a slope estimated separately for each year (including 1966), service class (first class or coach service), and aircraft type (propeller or jet) gives a R2 of 0.98 or higher in each regression, with an average R2 of 0.993.

Second, all airlines operating within the same route charged exactly the same price. In 1951, in our digitized price data we have 432 airport pairs in which both American Airlines and TWA were operating and reported the price for first class service. 94% of those airport pairs had exactly the same price in both airlines.

Third, ticket prices of flights operated by jet airplanes had a surcharge of around 6% on

[104] The sample of prices digitized was limited due to data availability.

top of the one operated by propeller airplanes.

Fourth, the change in prices over time had a similar pattern until 1961: a stronger increase in short distances (probably due to an increase in fixed costs of take-off and landing, although not reflected in the intercept of the linear regressions), and a relatively constant increase for flights between airports more than 1,000 km apart. In the period 1961 to 1966 we observe a drop in prices of around 20% for routes of more than 1,000km distance, breaking the linearity of prices in distance previously observed. We had visually inspected price tables and detected that the drop in prices happened in 1962-1963.

Figure 1.49 shows prices for first class service by year and aircraft type, deflated by the consumer price index to 1951 values. Figure 1.50 presents the percentage change in deflated prices of first class service. Both figures show the previous facts: prices are generally linear in distance until 1966 in which we observe a break after 1,000 km.

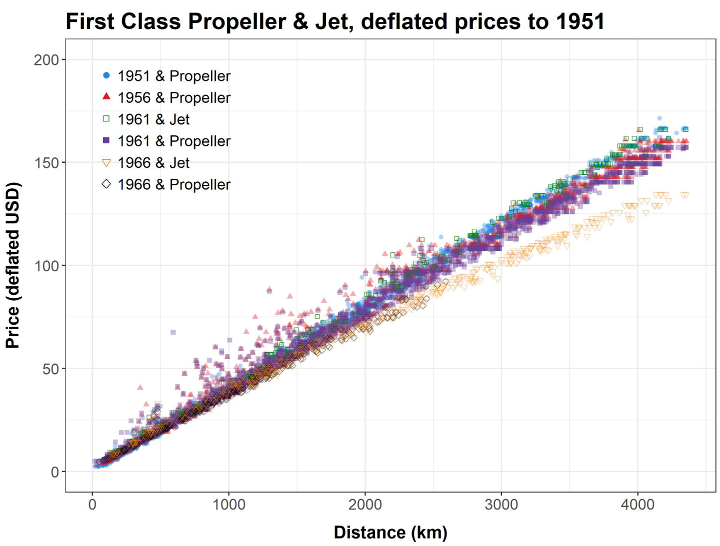


Fig. 1.49. Flight ticket prices, deflated by CPI

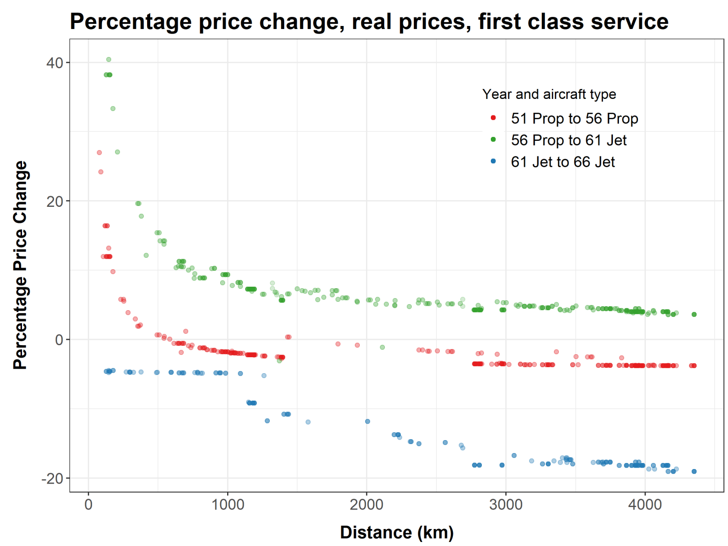


Fig. 1.50. Change flight ticket prices, deflated by CPI

We convert our sample of prices at the airport-pair level to prices of the population of MSA-pairs as follows: first, we obtain a pricing function that can flexibly approximate prices by regressing deflated prices on a cubic polynomial of distance separately for each year. We use prices of first class service for all years, propeller aircraft for 1951 and 1956 and jet aircraft for 1961 and 1966. Second, we predict prices for each MSA-pair and year using the MSA-pair distance and the year's estimated regression.

Highway travel time

Taylor Jaworski and Carl Kitchens have graciously shared with us data on county-to-county highway travel time and nominal travel costs for 1950, 1960 and 1970. Travel time is constructed using maximum speed limit in each highway segment and year. Travel

costs uses, for each year, travel time, highway distance, truck driver's wage and petrol costs. See Jaworski and Kitchens (2019) for details. The dataset is constructed using 2010 county boundaries and contains county centroids. We converted it to MSA-to-MSA by matching counties' centroids to 1950 MSAs using the shape file from Manson et al. (2020). We take the minimum travel time and minimum travel costs among all county pairs that belong to the same MSA pair. We convert nominal travel costs to 1950 real travel costs deflating by the consumer price index. We convert 1950, 1960 and 1970 travel times and travel costs to 1951, 1956, 1961 and 1966 by linearly interpolating (e.g. travel $\text{time}_{ij,1951} = \text{travel time}_{ij,1950} \times \frac{1960-1951}{10} + \text{travel time}_{ij,1960} \times \frac{1951-1950}{10}$).

The within MSA-pair correlation of the 1951-1966 change in travel time by highway and airplane is 0.068 for all MSA-pairs, and -0.011 for MSA-pairs more than 2,000 km apart. Figure 1.51 presents the MSA-pair 1951-1966 change in travel time by highway and airplane, where for exposition we only present MSA-pairs that had a reduction in travel time by both means of transport. Estimating a linear regression of change in air travel time on the change in highway travel time gives a slope of -0.02 not statistically different from zero, with a R2 of 0.00005.[105] Figure 1.52 repeats the exercise where MSA-pairs are weighted by the amount of establishment-technology pairs used to estimate the elasticity of citations to travel time (equation ((1.3))). In this case the estimated regression has a slope of 0.73 statistically significant at the 1% level and a R2 of 0.09.[106]

In Tables 1.5 and 1.26 we present the results of adding highway travel time as control. The low correlation between the change in travel time by highway and airplane implies that the estimated elasticity of citations to air travel time remains almost unchanged, relative to the baseline estimation.[107]

[105] 8.7% of MSA-pairs had an increase in travel time either by highway or by airplane. The regression with all MSA-pairs has a slope of 0.60 significant at the 1% level. However, the R2 of the regression remains very low: 0.0046.

[106] With all MSA-pairs the slope is 1.01 statistically significant at the 1% level and the R2 is 0.04.

[107] In order to perform a test of statistical difference of coefficients we would need to compute the covariance between the two regressions. Assuming the covariance is zero, in columns (1) and (2) 1.26 the coefficients of air travel time at +2,000km are not significantly different.

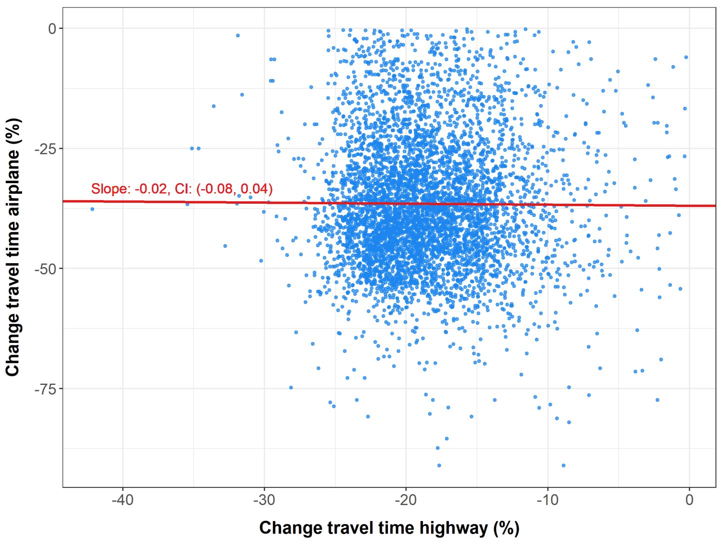


Fig. 1.51. Change travel time by airplane and highway 1951-1966

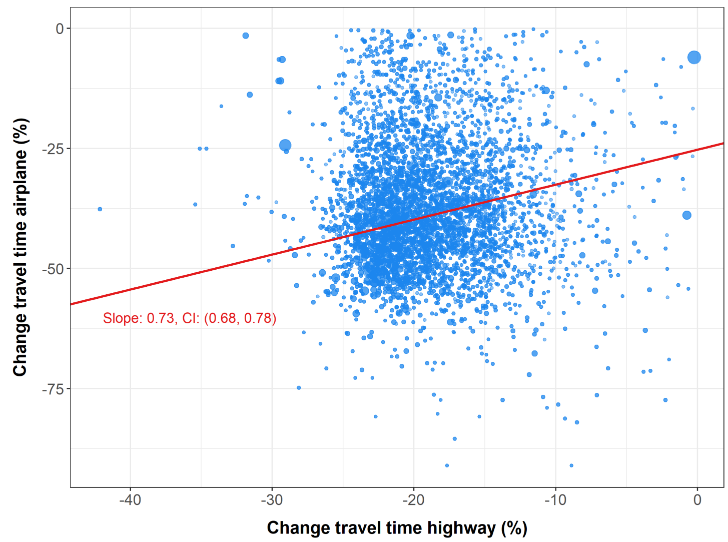


Fig. 1.52. Change travel time by airplane and highway 1951-1966, weighted

	PPML							
Dep. variable: *citations*	$cit_{FiGjhkt}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(travel time) $\times$ 0-300km	0.0213	0.0276	0.0198	0.0318	0.0252	0.0349	0.0283	0.0313
	(0.0388)	(0.0385)	(0.0391)	(0.0393)	(0.0389)	(0.0391)	(0.0396)	(0.0393)
log(travel time) $\times$ 300-1,000km	-0.0990***	-0.1040***	-0.0935***	-0.0745**	-0.1014***	-0.0857***	-0.0748**	-0.0861***
	(0.0269)	(0.0292)	(0.0265)	(0.0303)	(0.0290)	(0.0312)	(0.0303)	(0.0312)
log(travel time) $\times$ 1000-2,000km	-0.0928**	-0.1155**	-0.0710*	-0.0395	-0.0948*	-0.0498	-0.0318	-0.0435
	(0.0418)	(0.0485)	(0.0423)	(0.0523)	(0.0502)	(0.0573)	(0.0520)	(0.0576)
log(travel time) $\times$ +2,000km	-0.1848***	-0.1761***	-0.1724***	-0.1238**	-0.1658***	-0.1052*	-0.1236**	-0.1041*
	(0.0492)	(0.0531)	(0.0498)	(0.0587)	(0.0542)	(0.0607)	(0.0590)	(0.0609)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
Controls:								
log(highway time)	-	Yes	-	-	Yes	Yes	-	Yes
log(telephone share) $\times$ time	-	-	Yes	-	Yes	-	Yes	Yes
log(distance) $\times$ time	-	-	-	Yes	-	Yes	Yes	Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.26: Elasticity of citations to travel time: additional controls

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \mathbb{1}\{distance_{ij} \in d\}\, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \mathbb{1}\{distance_{ij} \in d\}\, \mathbb{1}\{X_{FiGjhkt}\}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0-300km]$, $(300km-1000km]$, $(1000km-2000km]$, $(2000km-max]$. Relative to (1), columns (2) to (8) contain additional controls. Log highway time between i and j changes in every time period t. The log mean share of households with telephone line in ij pair interacted in 1960 is interacted with a time dummy. Log distance ij is interacted with a time dummy. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Frequency adjusted travel time

The frequency of flights may have changed simultaneously with the introduction of jet airplanes. The change in travel time could then be consequence of higher frequency rather than changes in airplanes' speed. Given that some MSA pairs are connected indirectly (with connecting flights), accounting for frequency is not straight forward: the frequency of each leg of the flight route matters (actually, it is not only frequency of each leg but also the synchronization among all potential legs). In order to take into account potential changes in the frequency of flights we computed the daily average travel time. This travel time is the average across all fastest travel times if the passenger was to depart at each full hour (1am, 2am, ..., 1pm, 2pm, etc.). The computation of this travel time includes the

waiting time that is affected by frequency: the time until first departure and layover time of each connecting flight. Hence, the daily average travel time is a frequency-adjusted travel time: changes in the daily average travel time that are larger than in the fastest travel time denote that frequency of flights increased and therefore there is less waiting time. If we observe the reverse that means that frequency did not improve as much as the speed of airplanes.

Figure 1.53 shows the within MSA-pair decrease in the fastest travel time and the daily average travel time.[108] Both measures of travel time follow a similar pattern: slight decrease in 1956, a stronger decrease in 1961 especially for long distance routes, and a further decline in 1966. However, we observe that the decrease of the fastest travel time is on average larger than the one of the daily average travel time: the frequency of flights, if any, attenuated the potential decrease in travel time from the improvements in airplanes' speed. This observation is also in line with a comparison of the fastest travel time with and without layover time (Figure 28 in the Appendix of the paper): layover time attenuated the change in travel time.

In table 1.27 we estimated the elasticity of citations to travel time using first the fastest travel time (baseline, columns 1 and 2) and the daily average travel time (columns 3 and 4). The estimated elasticity is similar using both measures, which gives confidence that our results are not driven by changes in the frequency of flights.

[108]The within MSA-pair correlation of the (1951-1966) change in fastest travel time and the change daily average travel time is 0.60.

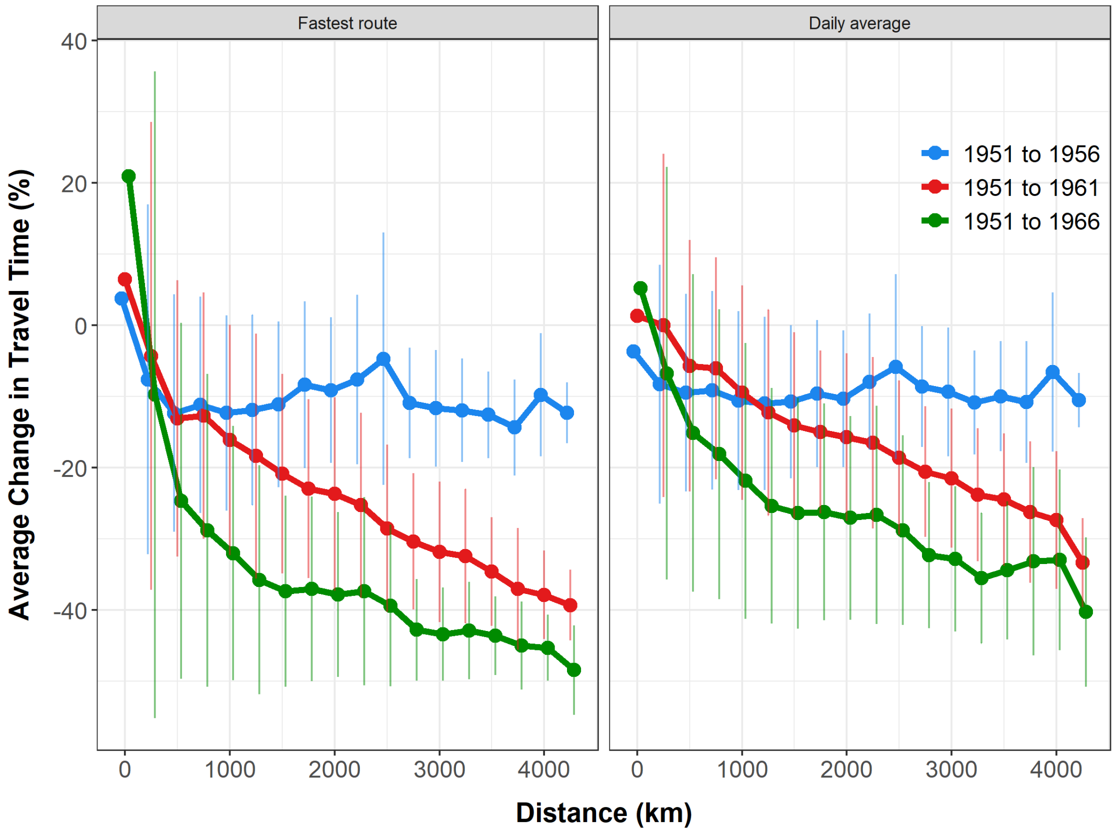


Fig. 1.53. Change in MSAs travel time: fastest travel time and daily average travel time

	PPML not bias-corrected			
Dep. variable: *citations*	$cit_{FiGjhkt}$			
	(1)	(2)	(3)	(4)
log(travel time)	-0.088^{***} (0.024)			
log(travel time) × 0-300km		0.021 (0.039)		
log(travel time) × 300-1,000km		-0.099^{**} (0.027)		
log(travel time) × 1000-2,000km		-0.093^{**} (0.042)		
log(travel time) × +2,000km		-0.185^{***} (0.049)		
log(travel time daily avg)			-0.100^{***} (0.039)	
log(travel time daily avg) × 0-300km				0.034 (0.037)
log(travel time daily avg) × 300-1,000km				-0.142^{***} (0.047)
log(travel time daily avg) × 1000-2,000km				-0.170^{***} (0.072)
log(travel time daily avg) × +2,000km				-0.236^{***} (0.064)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.27: Elasticity of citations to travel time: daily average travel time

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (2) includes the interaction of travel time$_{ijt}$ with a dummy for distance bin between the citing establishment Fi and the cited establishment Gj. Column (3) and (4) use the daily average travel time, which is computed as the average of the fastest travel time departing at every full hour (the average across all 24 potential departing times). Standard errors clustered at the non-directional location are presented between parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

E.2. *Creation of knowledge*

E.2.1. *IV PPML: first and second stage estimation*

	First stage OLS	Second stage PPML
Dep. variable:	log(knowledge access)	$Patentscit_{Fiht}$
	(1)	(2)
log(knowledge access fix routes)	1.01^{***} (0.032)	
log(knowledge access)		11.24^{*} (6.35)
residual		-2.31 (7.20)
N obs. effective	991,480	91,480
R2	0.99	0.85
Within R2	0.53	

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.28: Elasticity of patents to knowledge access: first and second stage IV PPML

	OLS First stage reference quartile	OLS First stage 3rd quartile	OLS First stage 2nd quartile	OLS First stage 3rd quartile	Second stage PPML
Dep. variable:			log(knowledge access)		$Patents_{Fiht}$
	(1)	(2)	(3)	(4)	(5)
log(knowledge access time fix routes)	1.00^{***} (0.03)	0.01 (0.06)	0.03 (0.03)	0.00 (0.01)	
log(knowledge access fix routes) × 3rd quartile	0.01^{*} (0.004)	1.11^{***} (0.03)	-0.00 (0.00)	-0.00 (0.00)	
log(knowledge access fix routes) × 2nd quartile	0.00 (0.01)	-0.01 (0.04)	1.11^{***} (0.03)	-0.00 (0.00)	
log(knowledge access fix routes) × 1st quartile	0.01 (0.01)	-0.00 (0.04)	-0.04 (0.04)	1.15^{***} (0.04)	
log(knowledge access)					10.26 (6.38)
log(knowledge access) × 3rd quartile					2.32^{***} (0.66)
log(knowledge access) × 2nd quartile					4.21^{***} (0.84)
log(knowledge access) × 1st quartile					5.77^{***} (1.11)
residual					-2.25 (7.27)
residual × 3rd quartile					-2.55 (1.59)
residual × 2nd quartile					-4.32^{**} (1.97)
residual × 1st quartile					-8.27^{**} (3.28)
N obs. effective	991,480	991,480	991,480	991,480	991,480
R2	1.00	1.00	1.00	1.00	0.85
Within R2	0.53	0.89	0.90	0.90	

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.29: Elasticity of patents to knowledge access: first and second stage IV PPML

E.2.2. *Robustness*

	Baseline	Quartile absolute	Quartile per capita
Dependent Variable: *Patents*		$Patents_{Fiht}$	
	(1)	(2)	(3)
log(knowledge access)	10.14*** (3.66)	9.36** (3.69)	7.77** (3.70)
log(knowledge access) × quartile 0.50		2.05*** (0.58)	0.75** (0.34)
log(knowledge access) × quartile 0.25		3.80*** (0.90)	1.58*** (0.50)
log(knowledge access) × quartile 0.00		5.00*** (1.30)	4.03*** (0.77)
N obs. effective	991,480	991,480	991,480
R2	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.30: Elasticity of new patents to knowledge access: absolute and per capita MSA innovativeness

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed within technology using the absolute level of patents in the MSA-technology in 1949-1953. Column (3) computes the quartile of innovativeness using patents per capita in the MSA-technology in 1949-1953 using 1950 population. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

	PPML		β by distance		+300km		+1,000km		+2,000km	
Dependent Variable: *Patents*					$Patents_{Fiht}$					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(knowledge access)	10.14*** (3.66)	9.36** (3.69)	18.17*** (4.63)	16.50** (4.76)	10.09** (4.66)	8.70* (4.67)	18.82*** (5.82)	19.08*** (5.74)	12.70 (8.18)	10.26 (7.92)
log(knowledge access) × 3rd quartile		2.05*** (0.58)		2.70*** (0.84)		2.12*** (0.58)		2.08*** (0.53)		1.94*** (0.49)
log(knowledge access) × 2nd quartile		3.80*** (0.90)		5.96*** (1.42)		4.19*** (0.88)		3.97*** (0.81)		3.64*** (0.73)
log(knowledge access) × 1st quartile		5.00*** (1.30)		8.94*** (1.97)		5.49*** (1.25)		5.28*** (1.23)		4.68*** (1.07)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.31: Elasticity of new patents to knowledge access, varying beta or distance.
Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Relative to columns (1) and (2), columns (3) and (4) compute Knowledge Access using four distance-specific β parameter according to distance bins between i and j. The bins are [0km, 300km], (300km, 1000km], (1000km, 2000km], +2,000km. Columns (5) to (10) use the same β as column (1) and (2), but computing Knowledge Access with a truncated sample of j that are further than a certain distance threshold from i. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

	PPML		OLS	
Dependent Variable: *Patents*	$Patents_{Fiht}$		$log(Patents_{Fiht})$	
	(1)	(2)	(3)	(4)
log(knowledge access)	10.14^{***} (3.66)	9.36^{**} (3.69)	6.83^{*} (3.19)	6.27^{*} (3.20)
log(knowledge access) $\times$ 3rd quartile		2.05^{***} (0.58)		0.92^{*} (0.51)
log(knowledge access) $\times$ 2nd quartile		3.80^{***} (0.90)		2.64^{**} (1.03)
log(knowledge access) $\times$ 1st quartile		5.00^{***} (1.30)		3.82^{**} (1.79)
N obs. effective	991,480	991,480	300,539	300,539
R2	0.85	0.85	0.87	0.87

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.32: Elasticity of new patents to knowledge access: PPML and OLS

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (3) estimates $log(\text{Patents})_{Fiht} = \boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht} + \xi_{Fiht}$. Columns (2) and (4) open the coefficient ρ by the quartile of innovativeness of location i within technology h, computed within technology using the absolute level of patents in the MSA-technology in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Difference in amount of observations is due to dropping zeros in columns (3) and (4). Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

Access to capital

We construct four measures of access to capital using 1949-1953 market capitalization of firms listed in the stock market. The four measures are similar in their essence but differ in the computation of a firm's technology and the firm's location. The measure is computed as follows:

$$capital\ access_{iht} = \sum_{k} \psi_{hk} \sum_{j,\, j\neq i} \text{Capital stock}_{jk,t=1951} \times \text{ travel time}^{\xi}_{ijt} \tag{1.15}$$

where $\text{Capital stock}_{jk,t=1951}$ is a proxy for the capital which is specific to technology k located in j at the initial time period 1951. ψ_{hk} is an input-output weight of capital flows and ξ is the elasticity of capital flows between to travel time. As a proxy for capital we use market capitalization of firms.

We construct four measures of $capital\ access_{iht}$ which differ on: (i) the way we define the allocation of the firm's capital to each location (either using all inventors' locations or only the assigned headquarters), and (ii) the way we allocate a firm's capital across technologies (using the share of a technology within the firm, or relative to the national share of that

technology). We use COMPUSTAT as our source of data for market capitalization.

We proceed as follows:

1. Use share's market price at closure calendar year multiplied by the number shares outstanding. We use the variables *prcc_c* and *csho* to maximize coverage of firms given that other variables have missing value for many firms.
2. Take the yearly average market capitalization to maximize coverage (many firms have missing in a certain year). This step potentially introduces measurement error due to changes in total stock market capitalization but allows us to increase the amount of firms included in the sample.
3. Determine a firm's MSA using patent inventor location. Two ways to determine the location, 1. only HQ location, 2. all locations where the firm had inventors applying for patents in 1949-1953
4. Determine the share of each technology firm's technology using patent technology. Two ways to determine the share oftechnology: 1. the share of each tech within firm + share within firm relative to national share
5. In the absence of data on a capital input-output weight, assume it is the same as the technology input-output weight, i.e. $\psi_{hk} = \omega_{hk}$
6. In the absence of data on the elasticity of capital flows to travel time assume $\xi = -1$

The four measures of access to capital are as follows:

1. Attribute all capital to headquarters and use the absolute share of each technology in the firm
2. Attribute all capital to headquarters and use the share of each technology in the firm relative to the national share
3. Attribute capital to establishments using their pat share and use the absolute share of each technology in the firm
4. Attribute capital to establishments using their pat share and use the share of each technology in the firm relative to the national share

Table 1.33 shows the results of estimating the elasticity of new patents to knowledge access while at the same time controlling for capital access.

Sensitivity to β

Indirectly connected MSAs

If the 1951 flight network was constructed in order to connect city pairs that would see future growth in citations, we can alleviate this endogeneity concern by focusing only on indirectly connected pairs.

Table 1.34 presents PPML regressions not bias-corrected. Columns (1) and (2) are the

Dependent Variable: *Patents*					$Patents_{Fiht}$				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(knowledge access)	10.14^{***} (3.66)					9.96^{**} (4.50)	11.29^{***} (4.32)	10.67^{**} (4.70)	12.90^{***} (4.43)
log(finance access hq)		0.54^{**} (0.26)				0.02 (0.30)			
log(finance access hq rel)			0.40 (0.25)				-0.14 (0.28)		
log(finance access est)				0.56^{*} (0.31)				-0.07 (0.39)	
log(finance access est rel)					0.31 (0.30)				-0.39 (0.38)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.33: Elasticity of new patents to knowledge access and finance access
Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \zeta_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) to (5) use as regressor the finance access of establishments in location i technology h and time period t, where the measure of finance access changes across columns. Columns (6) to (9) estimate the regression using both knowledge access and finance access. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

β	ρ	$\beta \times \rho$	Predicted yearly growth p.p.	Share yearly growth explained	Predicted yearly growth differential p.p.	Share yearly growth differential explained
-0.186	10.14	-1.89	3.47	0.78	1.1	0.21
-0.1	19.35	-1.94	3.5	0.78	1.07	0.2
-0.2	9.4	-1.88	3.47	0.78	1.1	0.21
-0.3	6.1	-1.83	3.45	0.77	1.14	0.22
-0.4	4.48	-1.79	3.44	0.77	1.16	0.22
-0.5	3.52	-1.76	3.44	0.77	1.19	0.23
-0.6	2.91	-1.74	3.45	0.77	1.2	0.23
-0.7	2.48	-1.73	3.47	0.78	1.22	0.23
-0.8	2.17	-1.73	3.5	0.78	1.22	0.23
-0.9	1.93	-1.73	3.52	0.79	1.24	0.24
-1	1.72	-1.72	3.51	0.79	1.28	0.24
-2	0.58	-1.16	2.8	0.63	1.55	0.3
-5	0.04	-0.19	1.19	0.27	3.65	0.7
-8	0.09	-0.76	8.22	1.84	6.96	1.33
-10	0.11	-1.08	15.16	3.4	8.19	1.56
-20	0.13	-2.63	69.8	15.65	21.66	4.14
-50	0.16	-8.22	531.34	119.16	219.49	41.94
-100	0.12	-12.33	5428.85	1217.49	2971.74	567.91

Table 1.35: Effect of knowledge access on new patents: varying the value of elasticity of knowledge diffusion

baseline regressions (all MSA-pairs), columns (3) and (4) drop MSA-pairs that are ever connected with one leg (a non-stop flight), and columns (5) and (6) drop MSA-pairs that are ever connected with one flight number. The difference between non-stop and one flight number is that one flight number could serve multiple MSAs by making intermediate stops.[109] The estimated coefficients are in the ballpark of the initial estimates, especially for +2,000km, providing evidence that it is reasonable to use the pre-existing network as the baseline to construct the instrument.

	PPML not bias-corrected					
Dep. variable: *citations*	$cit_{FiGjhkt}$					
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.088^{***} (0.024)		-0.202^{***} (0.051)		-0.241^{***} (0.061)	
log(travel time) × 0-300km		0.021 (0.039)		-0.237^{***} (0.116)		-0.410^{**} (0.165)
log(travel time) × 300-1,000km		-0.099^{**} (0.027)		-0.147^{*} (0.081)		-0.210^{**} (0.095)
log(travel time) × 1000-2,000km		-0.093^{**} (0.042)		-0.157^{*} (0.092)		-0.216^{**} (0.109)
log(travel time) × +2,000km		-0.185^{***} (0.049)		-0.297^{***} (0.085)		-0.242^{***} (0.090)
N obs. effective	4,703,010	4,703,010	1,735,427	1,735,427	1,396,393	1,396,393
R2	0.88	0.88	0.94	0.94	0.94	0.94
Observation selection:						
All	X	X				
Discard one leg			X	X		
Discard one flight number					X	X

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 1.34: Elasticity of citations to travel time: dropping directly connected MSA pairs

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (2) includes the interaction of travel time$_{ijt}$ with a dummy for distance bin between the citing establishment Fi and the cited establishment Gj. Column (3) and (4) discards all ij that are ever connected with one leg (non-stop flight), while columns (5) and (6) discard all ij that are ever connected with one flight number. The difference between non-stop and one flight number is that one flight number could serve multiple MSAs by making intermediate stops. Standard errors clustered at the non-directional location are presented between parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

[109] For example, in 1951 NYC-LA was connected with one flight number that included one stop in Chicago, that is two legs but only one flight number: passengers did not have to change airplanes).

Chapter 2

Time after Time: Communication Costs and Inventor Collaboration in the Multinational Firm

This chapter is co-authored with Çağatay Bircan (EBRD & UCL) and Beata Javorcik (EBRD & Oxford).

Abstract

We show that knowledge creation, as measured by patents, is increasingly conducted in cross-border collaborative teams of inventors. We document the importance of cross-border communication costs by showing that a higher overlap in business hours is associated with increased cross-border collaboration. This effect is distinct from the effect of physical distance, which matters as well. It is stronger for technology classes where lab experiments are involved and thus more frequent interactions may be required. Episodes of telecommunications liberalization (and the resulting decline in the cost of international calls) lead to an increase in cross-border collaboration, particularly when the business hour overlap between the headquarters and a subsidiary is larger. This effect is stronger for experiment-based technology classes. Less successful inventors respond more than their most successful peers.

JEL Classification: F14, F23, L23, O32, O33
Keywords: FDI; Patents; Collaboration; Knowledge Diffusion; Time Zones

1. Introduction

Knowledge creation and diffusion are the pillars of modern growth theory. Multinational enterprises (MNEs) play a central role in both the creation and diffusion of knowledge across international borders. According to conservative estimates of UNCTAD (2005), MNEs account for close to half of all global R&D expenditures and at least two-thirds of business R&D expenditures.[1]

Despite the importance of R&D efforts undertaken by MNEs, there is little micro evidence on this subject. Where do MNEs create knowledge? Is knowledge creation within a multinational firm becoming more concentrated in a few geographic locations, or is there more collaboration taking place across borders? What are the impediments to knowledge creation inside the boundaries of the firm?

Due to its very nature, knowledge creation is difficult to measure. We use data on patents and cross-border collaboration in inventor teams to capture the incidence of knowledge creation. We follow Kerr and Kerr (2018) in defining "global collaborative patents". These are innovations that involve at least one inventor located in the MNE home country and at least one inventor located in another country. Our data cover the 1980-2014 period and include geographic location of individual inventors within a team that obtained a patent. Matching patent data with the Orbis database, we are able to observe links between various establishments within a multinational firm.

In our analysis, we focus on three drivers of knowledge creation within an MNE: (i) time zone differences that lead to heterogeneity in the business hour overlap, (ii) episodes of telecom sector liberalization, which in the 1990s and the early 2000s resulted in a sudden and substantial decline in the price of international calls, and (iii) the interaction between the two.

Why should time zones matter for repeated interaction beyond the role of physical distance? Although much of physical production can be fragmented into individual stages and carried out relatively independently, innovative activity involves exchange of knowledge that is both tacit and strategic to firms. Sociological studies suggest that work practices in multinational organisations that involve knowledge work have evolved to demand

[1]The European Commission estimates that, in 2007, foreign-owned firms accounted for 15% of all business R&D in the United States; 20-25% in France, Germany, and Spain; 30%-50% in Canada, Hungary, Portugal, the Slovak Republic, Sweden, and the United Kingdom (UK); and more than 50% in Austria, Belgium, the Czech Republic, Malta, and Ireland (Dachs et al., 2012).

greater hours, commitment, and flexibility from their employees.[2] In economics, Chauvin et al. (2020) show that temporal distance stemming from time zone differences reduce synchronous and impromptu communication from first-best levels within a multinational organisation, presenting costly frictions especially for the multinational's knowledge-intensive work. Time zones differences have been shown to be a barrier to women sharing in the benefits of activities by firms engaged in international trade (Bøler et al., 2018).

The last few decades have witnessed a dramatic decline in the price of international calls (see Figure 2.10). This was largely driven by the liberalization of telecommunications markets, a process that started in 1984 in the US with the UK, Japan and New Zealand following suit shortly and the reforms gaining speed in Europe in the 1990s. Lower prices of international calls were accompanied by an increase in the volume of international calls, which in turn facilitated cross-border cooperation, including cooperation in R&D.

In this study, we hypothesize that a decline in communicating costs facilitated cross-border knowledge production between MNE establishments in home and host countries, particularly when the time zone differential was not too large. Moreover, we hypothesize that these factors were more important for experiment-based technology classes where more frequent communications between inventors may be required. Finally, we explore the heterogeneous impact on inventors with different patenting histories. On the one hand, inventors with lesser track records may stand to benefit more from international collaborations and thus may be willing to incur the high costs of international communication and the inconvenience of time zone differential. On the other hand, the presence of such costs, which must be partially born by their collaborators located in HQ, makes them less attractive as team members. Thus, it is ambiguous a priori whether inventors with a lesser or with a stronger track records would be affected more by a decline in communication costs.

In our core analysis, the outcome of interest is the share of patents produced by inventors in a particular foreign affiliate of an MNE that involved cooperation with one or more inventors from HQ. Our variables of interest include bilateral telecom liberalization between the MNE home and host country and the interaction of liberalization with the business hour overlap between the affiliate and the HQ. We define liberalization as both the home and the host country having liberalized their markets, or explicitly focus on the cost of international calls between the two countries. By considering the share of global

[2]For instance, Kvande (2009) discusses evidence from multinational law and computing firms that require employees to adjust working hours to collaborate with international business partners.

collaborative patents in all patents linked to a given foreign affiliate in a given time period, we are implicitly controlling for all factors that drive the volume of innovative activity in each location. By controlling for affiliate-HQ-technology fixed effects, we exploit the variation over time and thus abstract from the possibly endogenous decision of where to locate a foreign subsidiary. We also control for unobservable host-country-year hetorogeneity, and hence take into account factors that may have driven liberalization as well as variation in the supply of potential inventors that can be hired by an MNE in a given host country in a given year. Inclusion of these fixed effects means that our identification comes from comparing the impact of liberalization on affiliates located in the same host country but differing in terms of business hours overlap vis a vis their parent firm's HQ. Finally, technology-specific shocks are accounted for by technology-year fixed effects.

We extend the baseline analysis in three dimensions: (i) an event study focusing on telecom liberalization episodes and comparing affiliates with high versus low time zone differential vis a vis HQ; (ii) an examination of whether the size of inventor teams was affected by telecom liberalization and time zone differential; and (iii) an inventor-level analysis allowing for heterogeneous effects on inventors with different patenting histories.

Our econometric analysis produces three main sets of findings. First, in a simple cross-sectional setting, we find that a greater overlap in business hours between MNE HQ and affiliates leads to a higher incidence of inventor collaboration. For instance, an increase in business hour overlap by eight hours is associated with a 39% increase in the probability that a patent filed by inventors located in foreign affiliates involves HQ. In other words, an inventor working for a Polish subsidiary of a German MNE is 33% more likely than an inventor located in its Japanese subsidiary to collaborate on a patent with colleagues at the firm's HQ. The business hour overlap matters beyond the effect of physical distance, which itself reduces the likelihood of cross-border collaboration (by 8% when comparing a Japanese and a Polish subsidiary of a German MNE). It is also stronger for experimental technology classes, where more frequent interactions are likely to be required.

Second, we show that episodes of telecom sector liberalization have led to an increase in cross-border collaboration, particularly when the business hour overlap between the HQ and a subsidiary was larger. The results are robust to using an indicator variable for liberalization episodes or the cost of international calls between the two countries. Again we show that the impact is stronger for experiment-based technology classes. These conclusions are confirmed by event studies following the methodology by Borusyak et al.

(2021) and comparing the impact of telecom liberalization episodes on affiliate-HQ pairs with a large versus small business hour overlap. Let's stick with our example of a German MNE with a Polish and a Japanese subsidiary to illustrate the magnitudes. Our results suggest that after telecom liberalization, the share of patents based on collaboration with HQ filed by inventors located in the Polish subsidiary will increase by 6.4pp (22%). In contrast, collaborative patents in the Japanese subsidiary will go up by only 2.6pp (10%).

Third, one may expect to observe an increase in the inventor team size as the communication frictions decline. This is indeed what we find. We show that the number of inventors listed on a patent increases after telecom liberalization, particularly in affiliates having a larger business hours overlap with HQ. The effect is driven by inventors from HQs and not by inventors located in affiliates.

Finally, our inventor-level analysis confirms the earlier conclusions. It also indicates that the impact of telecom liberalization in affiliates with high business hour overlap vis a vis HQ is larger for inventors with lesser track records than for their colleagues. This is suggestive of high costs of working across time zones. It is consistent with the scenario where high monetary and inconvenience costs of working across time zones must be compensated by high expected benefits of collaboration and thus favor collaboration with inventor with a proven track record. Only when such costs decline, collaboration becomes worthwhile with inventors from foreign affiliates who have less of a track record.

We contribute to several strands of literature. First, we contribute evidence to the literature on where and how knowledge work is conducted within the multinational firm. Canonical models of foreign direct investment (FDI) posit a distance-concentration trade-off, which focus on trade in goods and do not take into account knowledge transfer (Helpman et al., 2004). Recent studies differ on how R&D and knowledge production are incorporated into models of FDI. Bilir and Morales (2020) model knowledge creation as concentrated in the HQ country and exploited abroad. In Keller and Yeaple (2013), a distance-knowledge trade-off emerges because it is more costly to transfer knowledge by direct communication than by trading intermediates. Similarly, Gumpert (2018) models how communication costs limit the ability of a firm's establishments to access knowledge at the headquarters.

Our findings support the existence of substantial knowledge transfer costs, both due to time zone differences and physical distance. They also show that multinationals increasingly conduct their R&D operations outside their home countries and in collaborative teams of inventors located in multiple countries. As such, the evidence is reminiscent of a vertical

model of FDI as in Antràs et al. (2006), who study the formation of cross-country teams in production. It is also in line with a theory of the multinational firm as an organization that specializes in the creation and transfer of knowledge across borders (Kogut and Zander, 1993).

Second, we contribute to the body of evidence documenting the importance of communication costs and time zone differences on multinational firm organization. Stein and Daude (2007) find that differences in time zones negatively affect FDI, while Oldenski (2012) finds that activities requiring complex within firm communication are more likely to occur at MNE HQ. Closest to our study is Bahar (2020), who presents evidence of a trade-off between distance to the HQ and knowledge intensity of the affiliates' industry. We extend this literature by explicitly showing that a decline in international communication costs has differential effects depending on the business hour overlap between the HQ and a foreign affiliate.[3]

Third, we add to the literature on cross-border collaboration in knowledge work. Kerr and Kerr (2018) find, in a sample of publicly listed companies from the United States (US), that global collaborative patents are frequently observed when a firm enters a new foreign region for innovative work, especially where intellectual property protection is weak. They also find that collaborative patents are higher quality than patents produced by inventor teams located only in the US, and employment of ethnic inventors at home is related to cross-border collaboration.[4] Catalini et al. (2020) show that travel costs constitute an important friction to collaboration between inventors, especially for high-quality scientists. Similarly, using the introduction of the jet engine to civil aviation in the 1950s as an exogenous reduction in travel time, Pauly and Stipanicic (2021) find that a decrease in travel time between two cities lead to an increase in patent citations between them. Our inventor-level analysis extends this literature by pointing out how the interaction between a drop in communication costs and the time zone differential affect collaborations of inventors with different track records. We show that a decline in communication frictions tends to benefit inventors with lesser track records more.

[3]One proposed reason for the negative relationship between distance (both physical and cultural) and FDI is the difficulty of a parent firm to monitor the activities of its affiliates abroad (Blonigen et al., 2020). Monitoring costs can be especially high in the context of innovative activity, where parent firms have an incentive to protect the leakage of proprietary technology. Our results suggest that business hour overlap may contribute to the difficulty of monitoring.

[4]Related, Foley and Kerr (2013) find that increases in the share of a US multinational's innovation performed by inventors of a particular ethnicity at home are associated with increases in the share of that firm's affiliate activity in countries related to that ethnicity.

Recent research suggests that ideas are getting harder to find (Bloom et al., 2020). Patents increasingly involve large research teams and evidence shows that interactions with better inventors are strongly correlated with subsequent productivity (Akcigit et al., 2018).[5] This increases the importance of collaboration across borders and ability of large teams of inventors to work together, often facilitated by within-firm mobility. Our work sheds new light on the determinants of cross-border collaboration.

Fourth, our paper adds to the literature on FDI and the geographic diffusion of knowledge and technology (Keller, 2004). Keller (2002) finds that productivity effects of R&D are declining in distance, while Bilir and Morales (2020) show that parent and affiliate R&D activities are complementary. We contribute to this literature by documenting that a decline in communication frictions lead to more R&D collaboration between HQ and MNE affiliates, thus increasing their complementarity.

The remainder of this paper is structured as follows. We describe our data in Section 2. The subsequent section presents stylised facts. Section 4 contains our results on the determinants of global collaborative patenting activity. Section 5 focuses on the telecom liberalization episodes at the level of the establishment, while in section 6 we disaggregate the analysis further down to the inventor level. Concluding remarks appear in Section 7.

2. Data

2.1. Patents

The main patent dataset underlying our analysis comes from USPTO's (United States Patent and Trademark Office) PatentsView project. PatensView covers the universe of US patents. Crucially, it contains inventor identifiers resulting from a disambiguation exercise and information on inventor location (at the city level). Inventor location allows us to pinpoint where knowledge creation takes place, while the inclusion of identifiers allows us to track inventors across time and space.

We combine the USPTO data with EPO's PATSTAT (Spring 2021 edition) using publication numbers. PATSTAT is an effort to collect data on patent filings from all over the world.

[5] Akcigit et al. (2018) introduce an endogenous growth model with knowledge diffusion in which inventors learn from each other via collaboration. They quantify the importance of interactions for growth by studying the effects of reducing interaction costs, such as IT or infrastructure, on inventors' learning and knowledge accumulation.

Importantly, this provides us with patent filings at EPO and JPO, two important patent offices other than the USPTO.

We focus on patent families, where a patent family is defined as a collection of patents concerning the same invention in potentially multiple patent offices around the world. In this way, we count a single invention only once and consider the date of its first filing as the relevant date. Our analysis focuses on triadic patent families, which include patents granted at all three of the patent offices mentioned earliers (USPTO, EPO, JPO).[6] These patent families capture the most important inventions relevant at a global level.[7]

2.2. *Building our data set*

The purpose of our analysis is to understand drivers of cross-border collaboration in innovation within a multinational firm. This requires several ingredients.

First, we need to be able to match patents to a firm. A firm can register legal ownership of a patent in a subsidiary that is located in a country different to the firm's HQ, different to the location where the underlying technology was created (i.e. where inventors reside), and different to the location where the intellectual property will be applied (Griffith et al., 2014). Therefore, it is crucial that we accurately identify the firm that is the ultimate owner of a patent and understand where exactly the invention was created. This requires assigning patents to firms. To do so, we use the Orbis Intellectual Property (IP) database, provided by Bureau van Dijk. Orbis IP sources its patents data from Lexis Nexis and maps the assignee (or patent owner) names indicated on the patent to Orbis firm identifiers based on a textual matching algorithm and extensive manual checks (see appendix A).

We then use Orbis's data on ownership links to identify the global ultimate owner (GUO) of the assignee. In the analysis below a GUO is our definition of a firm. The ownership links were extracted in September 2020 and they reflect the state of the world at that point in time. As we combine the data on ownership links with historical patent data, we may attribute some patents and inventors to a firm that at the time of the patent filing were not part of it, but that the firm subsequently acquired. We would then be confounding

[6]Note that triadic patents are sometimes defined as patent families granted in the US and filed, but not necessarily published, at EPO and JPO. We instead require the patent family to include a publication in all three offices. This ensures a reasonable sample coverage.

[7]Note that EPO uses two different definitions of patent families: the simple (DOCDB) and the extended (INPADOC) definition. We use the extended definition here which groups together all applications that have at least one priority in common. In what follows, we use the terms "patent" and "invention" interchangeably, both of which refer to the relevant patent family.

an effect that operates through acquisition with an effect operating in a fixed network of establishments. Our robustness check focusing only on foreign affiliates that already existed in the 1980s alleviates this concern.

Second, we need to know where the innovative activity actually took place. We obtain this information from the patent data, which give us the location of each inventor who contributed to the patent.

Third, we need to define what we mean by cross-border collaboration within an MNE. We focus on collaboration between inventors located in the MNE HQ country and those located in its foreign affiliates. While we could potentially consider collaboration between all pairs of entities within an MNE, this would result in a very large and sparse matrix. In addition, innovation is likely to follow a hub-and-spoke model, i.e., most collaboration takes place between an R&D hub in the home country and individual subsidiaries. We thus restrict our analysis to HQ-subsidiary collaboration.[8]

Fourth, we need to decide which MNE establishments to consider. Given the nature of the research question, we want to consider only establishments engaged in R&D, which in our case means establishments where at least one inventor listed on a patent belonging to an MNE was located. We define establishments based on patent data, using the information on the inventors' location of residence to identify the country and time zone where they are located.[9] Our definition of establishment comprises all inventors working for a given firm while located in a particular country and time zone. This implies that countries with a single time zone will have at most one establishment of a given firm, while countries with multiple time zones can have multiple establishments belonging to the same MNE. In other words, our approach amounts to pooling together patents for multiple establishments belonging to a single MNE and located in the same time zone of a single country. But it also implies that an MNE may have more than one establishment in its HQ country if the country has multiple time zones.

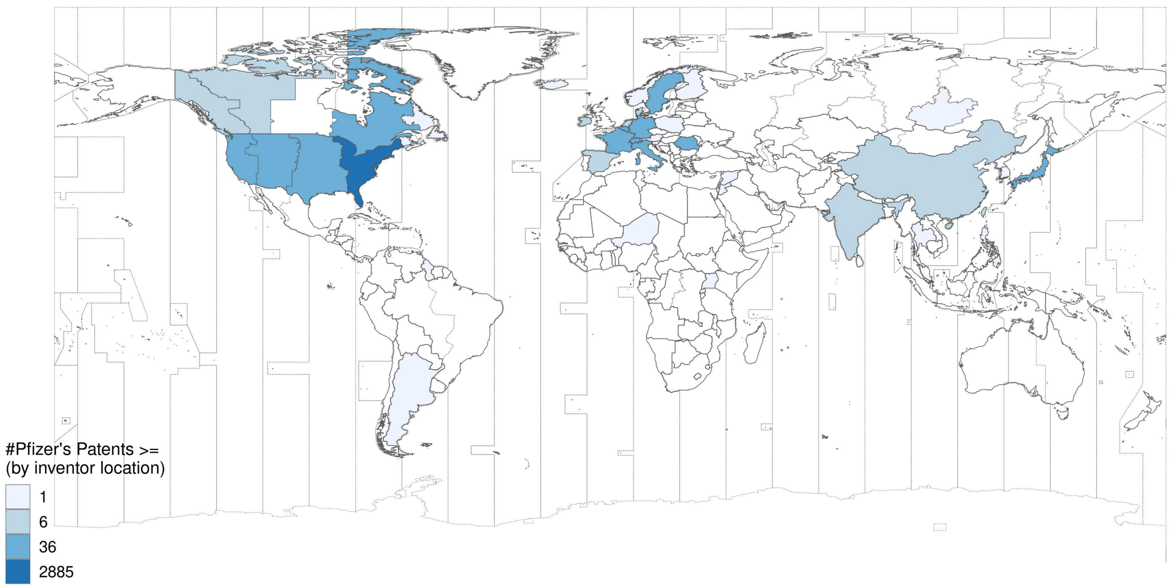
Figure 1 below illustrates our approach in the context of Pfizer. Pfizer's R&D HQ country is the United States, where the company has establishments in each of the four time zones.

[8] In our estimation sample of patents with at least one inventor based at a foreign affiliate, around 7.7% of patents are affiliate-affiliate collaborations without HQ-involvement. 29.8% of patents are in collaboration with HQ inventors (40.9% weighted by citations). Thus, collaboration indeed seems to follow a hub-and-spoke model, especially for the patents of highest quality. Nevertheless, affiliate-affiliate collaborations are also present in the data and should be studied further in future work.

[9] We use the R package lutz to infer the time zone from inventor locations.

Thus when we consider cross-border collaboration between the US and Canada within Pfizer, we will consider 12 cases of possible cooperation (4 Pfizer locations in the US x 3 Pfizer locations in Canada). We further consider collaborations between each of the 4 Pfizer locations in the US and all other countries where Pfizer has foreign affiliates.

Fig. 2.1. Pfizer's establishments



Our final dataset includes 509,884 patent families that are matched to a firm. It accounts for 83% of all granted triadic patents over the period 1980-2014 (or 84% when weighting by citations). For more details, see appendix Sample Coverage.

2.3. *Variable Definitions*

The following variables are used in our analysis.

- Collaboration: this is our outcome variable of interest. It captures the share of patents obtained by inventors located in a given establishment that involve at least one inventor based in the HQ country of the firm.
- Business Hour Overlap: We take the the difference in time zone between the HQ location and the foreign establishment location. The maximum difference is 12 hours. Then we define business hour overlap as 8 hours minus the time zone difference, setting negative values to 0, so that our overlap variable ranges from 0 to 8 hours.

- Distance: The establishment location is defined as the centroid of the country-time zone pair. The distance is the geodesic (or straight-line) distance between the two establishment locations.
- Technology: Our definition of a technology corresponds to a 4-digit IPC class. There are around 650 such classes.
- The year of liberalization of the market for international calls in a given country comes from Table 1 in Boylaud and Nicoletti (2000).
- Bilateral data on prices of international calls were collected from OECD reports for the years 1990, 1998 and 2003 (see OECD (1994), OECD (1999) and OECD (2003)). These contain bilateral rates for international phone calls from country A to country B at the peak hour. We convert these rates to a per-minute rate in 1990 US dollars. The rates pertaining to calls from A to B and from B to A are typically not the same, we consider the average or the minimum of the two in our analysis.
- Star inventors: We measure inventors' initial productivity by computing the number of all patents they filed prior to 1990, regardless of the firm filing the patent.[10] We then define as stars the top 25% of all inventors working in foreign establishments of a given firm. In a robustness exercise, we also use the top 10% or the (log) number of patents filed by an individual inventor before 1990. In yet another definition, we weight the number of patents by citations received before 1990 and recompute the top 25%.

2.4. Experiment-Intensive Technologies

In order to identify technologies that rely on high-intensity communication, we perform analysis on patent texts. We focus on the brief summary text provided by *PatentsView* for all patent applications filed at the USPTO in the year 2000 and search for the words "experiment" and "trial". We compute the share of patents containing at least one of these terms within each technology class and define a technology as experiment-intensive if this share is higher than 35.7% (the median share in the regression sample). For around 5% of observations, we are not able to classify the technology class using the procedure outlined above, as these technologies are not present in the sample of US patents we use for classification.

To get a sense of the broad fields in which our experiment-intensive technologies are concentrated, table 2.8 shows the share of experiment-intensive technologies (4-digit) within

[10]We capture any patent that appears in the USPTO's PatentsView data. This dataset contains patents granted from 1976 onwards.

each more aggregate technology grouping (1-digit). Among these technology groupings, chemistry patents involve the most experimentation and electricity and fixed constructions patents the least. Table 2.9 displays the 15 most common 4-digit technology classes in our sample and how they are classified. Again we see chemistry-related technology classes having the largest share of experimental patents.

In figure 2.9, we correlate the share of experimental patents by technology with other patent classifications. We show that experimental technologies tend to be more process-intensive (based on data from Ganglmair et al. (2022)) and more scientific (based on data from Marx and Fuegi (2020)). We will show below that experimental technologies nevertheless capture something distinct that matters for communication frictions.

3. Stylized Facts

This section contains a few stylized facts to motivate our analysis.

A large share of patenting activity takes place outside of MNE HQ countries and is driven by inventors located in foreign affiliates. Table 2.1 lists the top 10 HQ countries in our sample in terms of the number of patent families attributed to MNEs. The most innovative country in the sample, Japan, is the least collaborative: only around 11% of patents filed by Japanese MNEs involved also inventors located outside of Japan. In contrast, around 39% of all inventions filed by US MNEs involved at least one inventor from a foreign affiliate. European MNEs stand out when it comes to patenting innovations created outside HQ countries and the incidence of HQ-affiliate collaboration. In German MNEs, 28% of patents involved global collaboration. MNEs from the Netherlands and Switzerland have notably high levels of innovative activities abroad.

Table 2.1: Patent Families and Collaboration Patterns by Country

HQ Country	Number of patent families	Only HQ inventors (in %)	Only foreign affiliate inventors (in %)	HQ-affiliate collaboration (in %)
JP	110277	88.71	5.82	5.47
US	76772	60.60	16.33	23.06
DE	34077	58.63	12.92	28.45
FR	17353	50.81	24.91	24.28
KR	13555	77.29	6.54	16.17
GB	7031	34.79	32.20	33.01
SE	6384	48.29	26.10	25.61
IT	5311	55.71	13.22	31.07
NL	4796	8.42	48.81	42.76
CH	4571	5.73	55.83	38.44

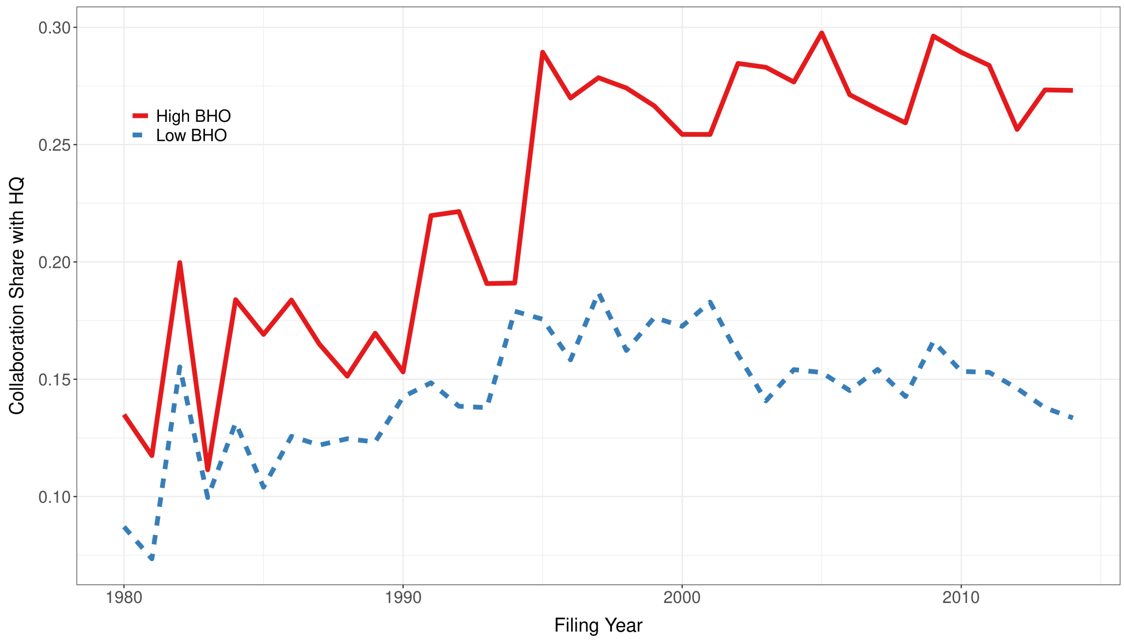
Liberalization of the telecom sector has led to a rapid decline in the cost of international calls and an exponential increase in the call volume. Table 2.10 presents the results of regressing call prices between country pairs on a binary variable that equals 1 if telecom sectors in both countries are liberalized. We find that telecom prices declined between 30% and 40% after liberalization[11]. Figure 2.11 shows the sharply increasing call volumes during the liberalization period[12].

Collaboration of affiliate-based inventors with HQ has been on the rise, especially in affiliates with a high overlap in business hours vis a vis the HQ. The share of cross-border collaboration in patenting has doubled from 13.5% in 1980 to around 27.3% in 2014 for establishment pairs with a high overlap in business hours (see figure 2.2). For affiliates located further away in terms of time zones, the increase in collaborations has been less impressive, as such collaborations went from from 8.7% in 1980 to 13.4% in 2014. Thus the temporal distance is alive and well, with the importance of time zone differences being exacerbated by improved communication technology.

[11] In figure 2.10 we plot a histogram of bilateral price changes between 1990 to 2003.

[12] We do not have data on the bilateral volume of calls and can thus not rerun the price regression with call volumes.

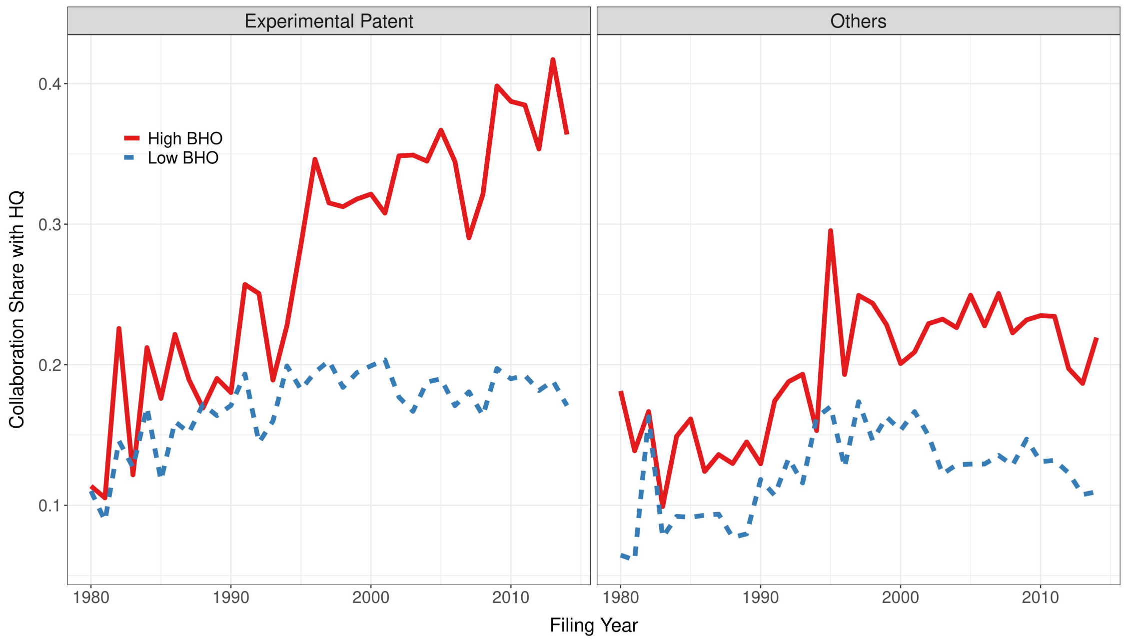
Fig. 2.2. Cross-border Collaboration over Time



Notes: This figure shows the share of patents filed by affiliate inventors that involve at least one HQ inventor. High overlap contains affiliates having more than 4 business hours overlap with HQ.

The rise in cross-border collaboration has been stronger in experimental technology classes. Inventors, located in affiliates with a high business hours overlap vis a vis MNE HQ, increased collaboration with HQ-based inventors, particularly in experimental technology classes. Their share of collaborative patents has tripled, increasing from 11.4% in 1980 to 36.4% in 2014. A much less pronounced rise has been registered in non-experimental technology classes or in affiliates with with a greater temporal distance to HQ (see figure 2.3).

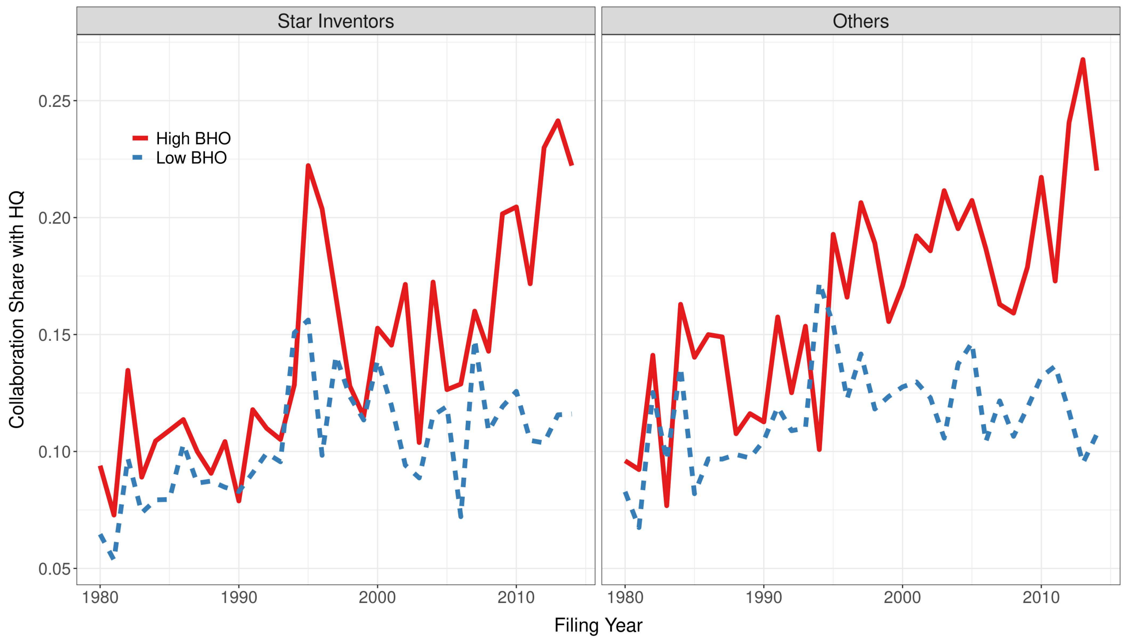
Fig. 2.3. Cross-border Collaboration over Time by Technology



Notes: This figure shows the share of patents filed by affiliate inventors that involve at least one HQ inventor. High overlap contains affiliates having more than 4 business hours overlap with HQ. Experimental technology classes are defined as in section 2.4.

The rise in cross-border collaborations has been driven by less experienced inventors in affiliates with greater business hour overlap with the HQ. Figure 2.4 shows a growing share of collaborative patents involving affiliate inventors with lesser temporal distance to HQs. This rise is particularly pronounced for non-star inventors, i.e., inventors with a lesser track record.

Fig. 2.4. Cross-border Collaboration over Time by Inventor Type



Notes: This figure shows the share of patents filed by affiliate inventors that involve at least one HQ inventor. High overlap contains affiliates having more than 4 business hours overlap with HQ. Star inventors are the top 25% inventors that had filed the most patents in USPTO data up to the year 1989 (see section 2 for details).

Patents filed by foreign-based inventors are of higher quality when collaborating with HQ. In table 2.11 we regress the number of forward citations on a binary variable which equals 1 if there is at least one HQ inventor on the patent. Even after controlling for firm, inventor, country-technology fixed effects and the (log) number of inventors collaborative patents receive around 10% more citations (significant at the 1% level). This motivates our focus on HQ collaboration.

4. Cross-border Collaboration and Time Zones

In the first pass at the data, we abstract from the telecom liberalization and examine the drivers of cross-border collaboration in a cross-sectional setting. We estimate the following equation:

$$\text{Collaboration}_{aht} = \alpha_1 \text{Overlap}_{ah} + \alpha_1 \text{Overlap}_{ah} \times \text{Experimental}_t + \text{FE}_{c(a)t} + \text{FE}_{f(a)} + \epsilon_{aht} \tag{2.1}$$

where $\text{Collaboration}_{aht}$ is the share of patents filed by inventors in foreign affiliate a in technology t that are in collaboration with inventors located in HQ establishment h of the same MNE.[13] Focusing on the share of patents as our outcome variable implicitly controls for all factors that may determine the scale of R&D activity in affiliate a or the likelihood of R&D outputs being converted into patents.

The explanatory variables of interest are the overlap in business hours between the time zone in which affiliate a is located and the time zone of the MNE's HQ establishment h (Overlap_{ah}) and its interaction with an indicator for experimental technology classes. The affiliate country-technology fixed effect captures unobservable heterogeneity related to innovative prowess of the host country. Thus our identification relies on variation in business hour overlap between foreign affiliates operating in country c and technology t and their parent companies in various home countries. The MNE fixed effect captures the possibility of different firms having different willingness to engage in cross-border collaboration in innovative activity. Standard errors are clustered at the location pair-technology level.[14]

We hypothesize that a high overlap in business hours reduces communication frictions and thus increases collaboration ($\alpha_1 > 1$). In addition, we expect technology classes that require a high frequency of interactions to be more sensitive to time zone differences. We thus interact our main independent variable with an indicator variable that captures whether t is an experimental technology class (Experimental_t) and expect $\alpha_2 > 0$.

The results presented in Table 2.2 confirm our hypothesis that business hour overlap facilitates cross-border collaboration. In all specifications, the coefficient on the business hour overlap is positive and statistically significant at the 1 percent level. Inventors in experimental technology classes benefit more from higher overlap in business hours, with the effect statistically significant at the 1 and the 5 percent level in columns 2 and 4,

[13] In the case of countries spanning multiple time zones, establishments located in different time zones will enter the sample as separate observations.

[14] The results are robust to clustering at the firm level.

respectively.

The magnitude of the estimated effects is economically meaningful. Column 1 implies that an increase in business hour overlap by eight hours is associated with a 39% increase in the probability of a patent involving HQ for inventors located at foreign affiliates. In other words, an inventor working for a Polish subsidiary of a German multinational is 39% more likely than an inventor located in its Japanese subsidiary to collaborate on a patent with colleagues at the firm's HQ.

There are several threats to our interpretation of the results. First, there may be other variables correlated with business hour overlap. In particular, geographical distance may be both correlated with temporal distance and important for collaboration patterns, e.g. because of increased travel time. In columns 3 and 4, we show that controlling for geographical distance barely changes our estimated coefficients. As previous studies we find a negative effect of geographical distance on collaboration. While an important determinant of collaboration, we find that physical distance matters less than temporal distance. Inventors based in Tokyo collaborate 8% less with their Berlin HQ than their colleagues in Warsaw. Finally, we find no evidence that the effect of physical distance matters more for experimental technologies, suggesting that business hours overlap captures a distinct friction to communication especially relevant for projects with a high frequency of interactions.

A second concern is that experimental technologies might feature other characteristics that make collaboration difficult when temporal distance is high. In a robustness check (not shown to save space), we classify technology classes by how much they rely on scientific articles and by how process- rather than product-intensive they are. Controlling for the interaction of these technology characteristics and business hour overlap barely affects the baseline results. In addition, using a continuous measure of the share of experimental patents by technology class yields a highly significant effect.

The cross-sectional analysis presented thus far provides suggestive evidence of a substantial effect of business hour overlap on collaboration patterns. Nonetheless, a causal interpretation of the results remains difficult. Firms may anticipate the difficulty caused by temporal distance and adapt the nature of investments they make. For example, a German MNC may strategically locate tasks or personnel in Japan requiring little collaboration whereas placing those with a high need for interaction in its Polish subsidiary. The following section addresses this concern by exploiting exogenous variation in communication

costs stemming from the liberalization of telecommunication sectors. By relying on temporal variation with affiliate-HQ-technology cell, it abstracts from concerns about location of new subsidiaries responding to telecom liberalization.

Table 2.2: Collaboration, Business Hour Overlap and Distance

	Share Collaboration			
	(1)	(2)	(3)	(4)
Overlap	0.0646***	0.0593***	0.0582***	0.0526***
	(0.0019)	(0.0020)	(0.0033)	(0.0034)
× Experimental		0.0111***		0.0120**
		(0.0028)		(0.0059)
log(distance)			-0.0100***	-0.0103***
			(0.0034)	(0.0035)
× Experimental				0.0014
				(0.0050)
Observations	316,875	316,875	316,875	316,875
R^2	0.47	0.47	0.47	0.47
Dependent variable mean	0.36	0.36	0.36	0.36
Fixed Effects				
GUO	✓	✓	✓	✓
Host Country-Technology	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.1. Standard errors are clustered at the location pair-technology level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

5. Liberalization of Telecommunication Markets

The next step in our analysis takes advantage of the temporal variation in the timing of telecom liberalization episodes across countries. The underlying intuition is that liberalization of a telecom market brings competition and leads to a decline in the price of international calls, which in turn boosts communications between affiliates and HQ and facilitates cross-border cooperation. We hypothesize that the impact of liberalization will be increasing in the business hour overlap between the affiliate and HQ. Put differently, we conjecture that if a large time difference makes communications very inconvenient, a drop

in the price of international calls is not going to compensate fully for this inconvenience.

Table 2.13 shows the year of telecom liberalization for OECD countries. Among them, only the US, UK and Japan liberalized in the 1980s, followed by other Commonwealth countries and Sweden in the early 1990s and Western Europe in 1998. Our empirical approach exploits this staggered pattern of liberalization.

5.1. Baseline

To test our hypothesis we estimate the following equation:

$$\text{Collaboration}_{ahty} = \beta_1 \times \text{Liberalization}_{c(a)c(h)y} + \beta_2 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Overlap}_{ah} + \\ + \text{FE}_{aht} + \text{FE}_{c(a)y} + \text{FE}_{ty} + \epsilon_{ahty} \quad (2.2)$$

where $\text{Collaboration}_{ahty}$ is the share of patents filed by inventors in foreign affiliate a in technology t in year y that are in collaboration with inventors located in HQ establishment h of the same MNE. $\text{Liberalization}_{c(a)c(h)y}$ measures whether markets for international calls have been liberalized in both the affiliate country $c(a)$ and the HQ country $c(h)$ in year y. Because of the so called termination fees, the liberalization status of both markets matters. To study whether the impact of liberalization varies with the type of communication required, we estimate equation 2.2 for the full sample, and separately for experimental and other technology classes.

We include several sets of fixed effects. First, all our specifications include establishment-pair-technology fixed effects (aht), meaning a fixed effect for firm f's affiliates in a particular time zone of host country $c(a)$ and a particular time zone of home country $c(h)$. This implicitly requires firm f's presence in a particular time zone of country c both before and after the two-sided liberalization of the telecom markets, thus eliminating the possibility that our results are driven purely by entry of MNEs into a new location. This also eliminates any other time-invariant factors that may drive research collaboration between two establishments, such as a common language or historical ties.

Second, we control alternatively for affiliate-country-time ($c(a)y$) or HQ-country-time ($c(h)y$) fixed effects. The former allow us to take into account unobservable affiliate-

country-time heterogeneity (that may be driving affiliate country liberalization) and use variation in geographic location of HQ to identify the effects of interest. Thus differences in timing of liberalization across pairs of countries and differences in time zones across host countries are the source of identifying variation. Alternatively, we account for unobservable HQ-country-time heterogeneity and use variation across affiliate countries of MNEs.

Third, we include technology-time fixed effects (*ty*), capturing any technology-specific changes in collaboration patterns. For example, as industries mature and products become more complex, innovation may require increasingly larger teams.

The results, presented in Table 2.3 are supportive of our hypothesis. The impact of telecom liberalization on cross-border R&D collaboration increases in the business hour overlap between the affiliate and the HQ. In most specifications, no statistically significant impact is present if there is no overlap in business hours. In addition to the continuous variable capturing overlap in business hours, we estimate a variation of equation 2.2, dividing the liberalization variable into two-hour intervals of business hour overlap (columns 4 to 6). We find that telecom liberalization has its biggest impact when temporal distance is lowest, i.e. a business hour overlap of at least 6 hours.

The liberalization impact is generally larger and varies more with business hour overlap for experimental technology classes, confirming our earlier results that communication frictions are more relevant in technology classes that plausibly require more exchange between inventors.

Table 2.3: Telecom Liberalization and Collaboration

	Share Collaboration					
	Full Sample	Experimental	Non-experim.	Full Sample	Experimental	Non-experim.
	(1)	(2)	(3)	(4)	(5)	(6)
Liberalization	0.0154	0.0234	0.0050			
	(0.0143)	(0.0167)	(0.0209)			
× Overlap	0.0208***	0.0258***	0.0124**			
	(0.0059)	(0.0081)	(0.0052)			
× 0-2 hours				0.0260*	0.0384**	0.0081
				(0.0141)	(0.0149)	(0.0212)
× 2-4 hours				0.0422	0.0372	0.0548**
				(0.0345)	(0.0474)	(0.0213)
× 4-6 hours				0.0446**	0.0452*	0.0561**
				(0.0201)	(0.0251)	(0.0236)
× 6-8 hours				0.0639***	0.0858***	0.0306*
				(0.0140)	(0.0176)	(0.0182)
Observations	575,780	293,241	282,539	575,780	293,241	282,539
R^2	0.84	0.81	0.87	0.84	0.81	0.87
Dependent var. mean	0.27	0.28	0.26	0.27	0.28	0.26
Fixed Effects						
Est. Pair-Technology	✓	✓	✓	✓	✓	✓
Year-Host Country	✓	✓	✓	✓	✓	✓
Year-Technology	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.2. Standard errors are clustered at the country-pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

To better understand the magnitude of the result, take the estimates in column 4 of table 2.3 and consider a German MNE with a Polish and a Japanese subsidiary with 8 and 0 hour business hour overlap, respectively. After telecom liberalization, the share of patents in collaboration with HQ filed by inventors located in the Polish subsidiary will increase by 6.4pp (22%). In contract, collaborative patents in the Japanese subsidiary will increase by 2.6pp (10%).[15] Thus, the effect of telecom liberalization on collaboration is more than twice as large for the Polish affiliate, both in absolute and relative terms.

Table 2.14 shows that the result is robust to focusing on affiliates that already existed in the 1980s. This further alleviatates concerncs about telecom liberalization triggering FDI, thereby leading to creation of new establishments. Table 2.15 in the appendix shows that results are robust to using the variation coming from affiliate-country liberalization

[15]The relative effect is computed using the pre-treatment mean for each interval of business hour overlap.

by including HQ-country-year fixed effects. Finally, instead of splitting the sample, we interact our liberalization and overlap variables with the indicator capturing experimental technologies. Table 2.16 shows that this triple interaction term yields a statistically significant coefficient, again supporting our conjecture that experimental technologies are affected more.

5.2. Event Study

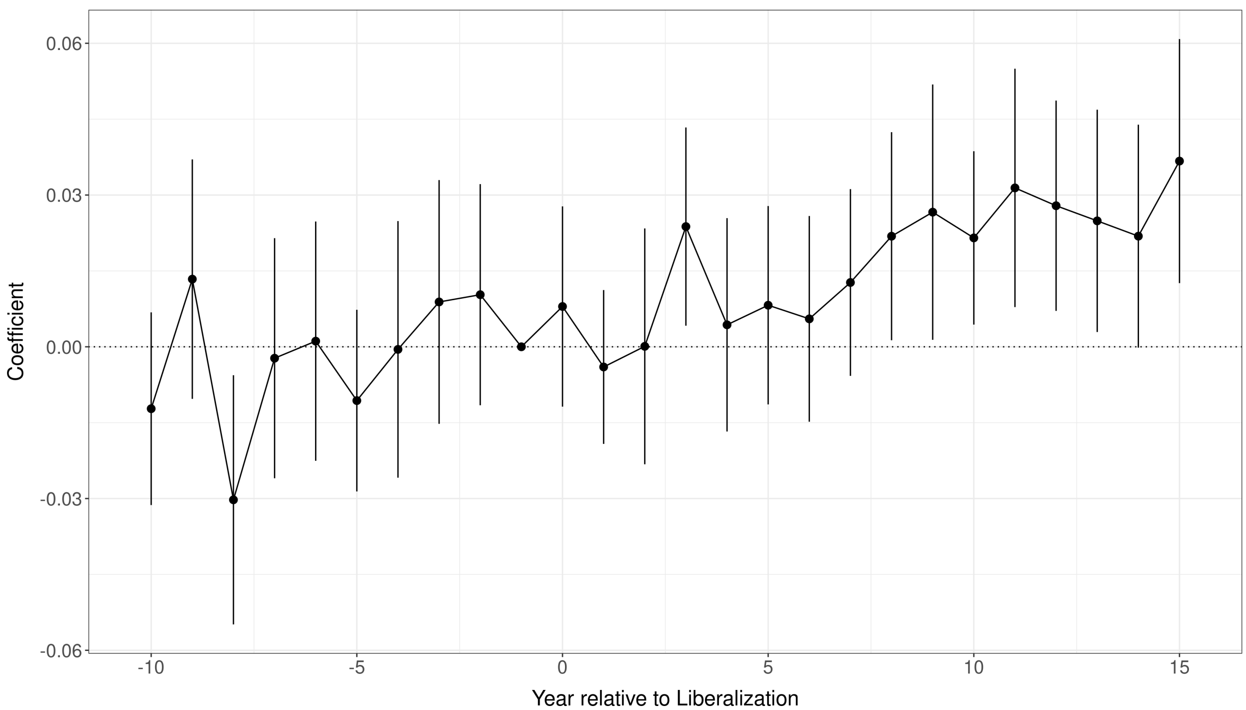
The panel nature of our data allow us to conduct event studies. Specifically, we test whether collaboration patterns were already on the rise prior to telecom liberalization in country pairs with high business hour overlap relative to those with low business hour overlap. This test also allows us to document how telecom liberalization affects collaboration over time. For the event study, we estimate:

$$\begin{aligned}\text{Collaboration}_{ahty} = &\sum_{\substack{d=-10 \\ d\neq -1}}^{d=15} \beta_{1,d}\mathbb{1}\{y = d + y_{0c(a)c(h)}\}+ \\ &+ \sum_{\substack{d=-10 \\ d\neq -1}}^{d=15} \beta_{2,d}\mathbb{1}\{y = d + y_{0c(a)c(h)}\} \times \text{Overlap}_{ah} + \text{FE}_{aht} + \text{FE}_{c(a)y} + \text{FE}_{ty} + \epsilon_{ahty}\end{aligned} \tag{2.3}$$

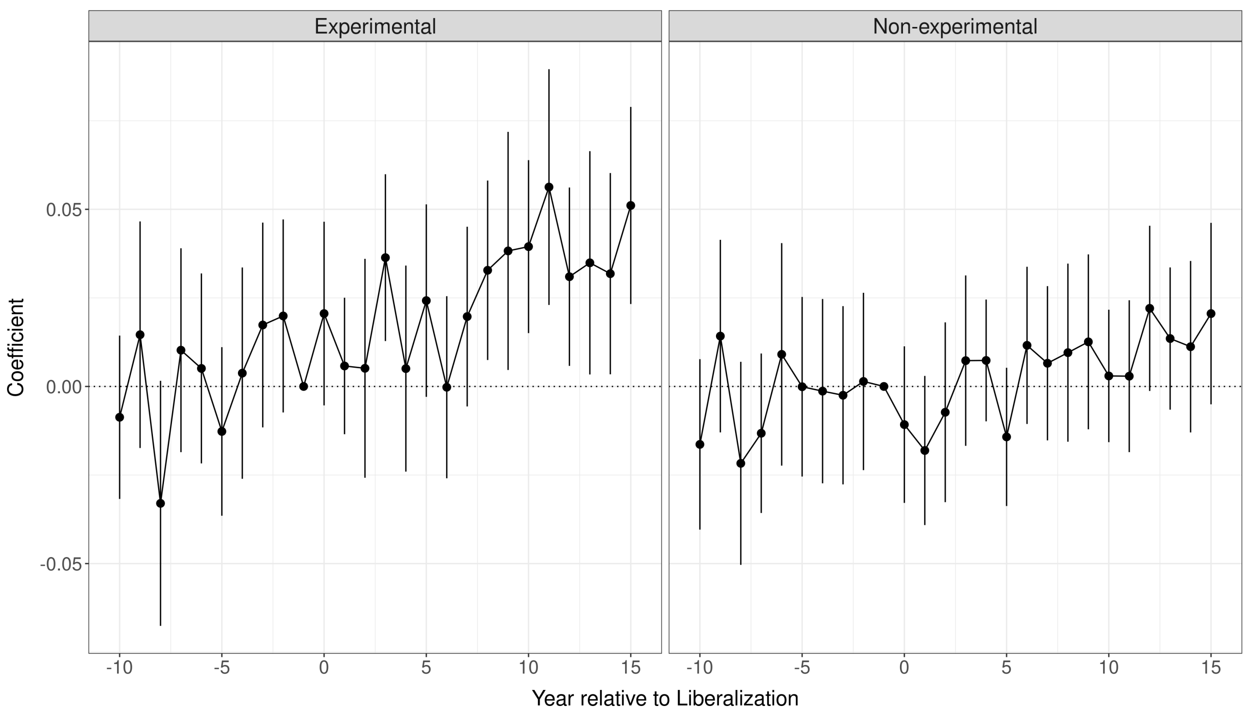
This specification resembles equation 2.2, but we allow for a time-varying impact of liberalization rather than relying on a simple pre- versus post-liberalization comparison. $y_{0c(a)c(h)}$ denotes the first year where both the affiliate country and HQ country have liberalized their telecom markets. Again, we estimate this model on the full sample and the two subsamples of technology types separately.

Figure 2.5 shows our estimates for β_2. We do not detect any changing collaboration patterns by business hour overlap prior to telecom liberalization. This lends support to the assumption that, in the absence of telecom liberalization, cross-border collaboration would have followed similar patterns in all country pairs regardless of business hour overlap. Panel b shows that the effect is more pronounced for experimental technologies. It takes around five years for the effect to kick in. This is to be expected as it takes time to start new collaborations and develop new patentable ideas.

Fig. 2.5. Event Study of the Impact of Telecom Liberalization on Collaboration



(a) Full Sample



(b) Split by Technology

Notes: This figure reports the results of estimating equation 2.3. Standard errors are clustered at the country-pair level. 90% confidence intervals are shown.

A recent set of methodological papers on difference-in-differences shows that traditional two-way fixed effects may produce misleading estimates when the treatment effect is heterogeneous between groups or over time.[16] Although these studies propose alternative estimators that are robust to such heterogeneity, they do not deal with cases as in our baseline equation (2.2), where the treatment variable, $Liberalization_{c(a)c(h)y}$, is interacted with a continuous term, $Overlap_{ah}$.

We therefore modify our event study specification in equation (2.3) to present results from one of these alternative estimators. Specifically, we redefine our treatment variable to equal 1 for affiliate and HQ country pairs that both liberalize their telecom markets *and* that have a business hour overlap of greater than four hours. We then employ the imputation estimator of Borusyak et al. (2021) to estimate the treatment effect in the ten years before and fifteen years after the liberalization episode.[17] Figure 2.12 in the Appendix confirms the lack of pre-trend effects prior to liberalization and corroborates the finding that cross-border collaboration has increased by around 5 percentage points in the decade following it (panel a). As in our baseline estimates, these effects are stronger for experimental technology classes (panel b) and statistically indistinguishable from zero for non-experimental technologies (panel c).

5.3. *International Call Prices*

An alternative (and a more direct) way of capturing telecom price liberalization is to focus on the actual price of international calls. Figure 2.10 shows a histogram of call rate changes from 1990 to 2003. As visible in the figure, prices of international calls declined over this period for essentially all country pairs, though with substantial heterogeneity. The average rate declined by around 160%. Whereas the US generally saw a low decline in international call rates over this period, the European liberalization of telecommunication markets in the late 1990s resulted in a large drop in rates for those countries.

Thus next, we exploit heterogeneity in changes of international call prices across country pairs and estimate the following equation:

[16] See Roth et al. (2022) for an overview of this literature.

[17] To achieve the imputation estimates, we can only include unit (affilitate and HQ country pair) and time (filing year) fixed effects in this revised specification.

$$\text{Collaboration}_{ahty} = \beta_1 \times \log(\text{Call Price}_{c(a)c(h)y}) + \beta_2 \times \log(\text{Call Price}_{c(a)c(h)y}) \times \text{Overlap}_{ah} + \\ + \text{FE}_{aht} + \text{FE}_{c(a)y} + \text{FE}_{ty} + \epsilon_{ahty} \quad (2.4)$$

This equation is very similar to our previous specification in equation 2.2, but there are two differences. First, the indicator variable for the liberalization episodes has been replaced with Call $\text{Price}_{c(a)c(h)y}$. This variable is defined as the per-minute price for an international phone call between the affiliate and the MNE's HQ country.[18] Second, because the data on call prices are available only for 1990, 1998 and 2003, y now represents 5-year periods (1990-1994, 1998-2002 and 2003-2007) instead of annual observations.

The results, presented in table 2.4 below, provide further support for our hypothesis. A decline in call prices, i.e. communications costs, boosts the share of cross-border collaborative patents only in the presence of some overlap in the business hours between HQ and affiliates. The mean reduction in minimum call prices between 1990 and 2003 in the regression sample is around 2.19 log points. This would translate to an increase in collaboration of around 7.8 percentage points for the German-owned Polish subsidiary and no statistically significant effect for the Japanese affiliate of the same firm.

[18]As call prices are not necessarily symmetric, i.e. calling from country A to B can be less/more expensive than vice versa, we take the minimum of both prices.

Table 2.4: International Call Prices and Collaboration

	Share Collaboration					
	Full Sample	Experimental	Non-experim.	Full Sample	Experimental	Non-experim.
	(1)	(2)	(3)	(4)	(5)	(6)
log(Minimum Call Price)	0.0008	-0.0055	0.0101			
	(0.0124)	(0.0149)	(0.0120)			
$\times$ Overlap	-0.0131***	-0.0136***	-0.0127***			
	(0.0027)	(0.0037)	(0.0037)			
$\times$ 0-2 hours				-0.0079	-0.0152	0.0022
				(0.0127)	(0.0149)	(0.0119)
$\times$ 2-4 hours				-0.0268*	-0.0118	-0.0549***
				(0.0154)	(0.0181)	(0.0192)
$\times$ 4-6 hours				-0.0306	-0.0448	-0.0153
				(0.0198)	(0.0436)	(0.0318)
$\times$ 6-8 hours				-0.0355***	-0.0430**	-0.0255**
				(0.0131)	(0.0167)	(0.0117)
Observations	232,577	114,138	118,439	232,577	114,138	118,439
R^2	0.95	0.94	0.96	0.95	0.94	0.96
Dependent var. mean	0.31	0.33	0.30	0.31	0.33	0.30
Fixed Effects						
Est. Pair-Technology	✓	✓	✓	✓	✓	✓
Period-Host Country	✓	✓	✓	✓	✓	✓
Period-Technology	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.4. Standard errors are clustered at the country-pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

5.4. *Team Size*

A decline in communication costs may be expected to an increase in the size of inventor teams. This is the question we turn to in this subsection and estimate the following specification:

$$\log(\text{Team Size})_{ahty} = \beta_1 \times \text{Liberalization}_{c(a)c(h)y} + \beta_2 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Overlap}_{ah} + \\ + \text{FE}_{aht} + \text{FE}_{c(a)y} + \text{FE}_{ty} + \epsilon_{ahty} \quad (2.5)$$

where the dependent variable now captures the average number of inventors listed on all collaborative patents involving foreign affiliate a and HQ establishment h in technology

class t in year y. The right-hand-side of the equation remains as before. We take logs of the dependent variable, since the size of inventor teams is right-skewed[19].

The results, presented in column 1 of table 2.5 mirror our earlier findings and confirm our priors. As anticipated, telecom liberalization leads to an increase in the size of inventor teams but only when there is a substantial business hour overlap between a foreign affiliate and HQ. To unpack these results, we disaggregate the dependent variable to measure the number of inventors involved based in a foreign affiliate (column 2) and in HQ (column 3). Greater response of HQ-based inventors seems to be driving the results.

Table 2.5: Liberalization and Team Size

	log(Team Size)		
Inventors Located in	HQ or Affiliate	Affiliate	HQ
	(1)	(2)	(3)
Liberalization	-0.0216	-0.0271	0.0056
	(0.0212)	(0.0229)	(0.0257)
$\times$ Overlap	0.0226**	-0.0039	0.0358***
	(0.0091)	(0.0099)	(0.0082)
Observations	575,780	575,780	575,780
R^2	0.79	0.73	0.88
Dependent variable mean	0.94	0.62	0.40
Fixed Effects			
Establishment Pair-Technology	✓	✓	✓
Year-Host Country	✓	✓	✓
Technology-Year	✓	✓	✓

Notes: This table reports the results of estimating equation 2.5. Standard errors are clustered at the country-pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

6. Inventor-level Analysis

So far, the focal point of our analysis has been a foreign affiliate of an MNE. In a final step of our study, we further exploit the granularity of patent data by conducting inventor-level

[19] When we focus on HQ inventors only, we use a $\log(1 + x)$ transformation instead, since the HQ team size can be zero.

analyses. This allows us to study whether the effects of communication frictions vary with inventor characteristics. In particular, guided by the existing literature, we distinguish between established or star inventors and their peers with a lesser track record.

6.1. Baseline

We start by estimating the following equation:

$$\begin{aligned}\text{Collaboration}_{iahy} = &\beta_1 \times \text{Liberalization}_{c(a)c(h)y} + \\ &+\beta_2 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Overlap}_{ah} + \\ &+\text{FE}_{ah} + \text{FE}_{c(a)y} + \text{FE}_i + \epsilon_{iahy}\end{aligned} \tag{2.6}$$

This equation is similar to our baseline establishment-level specification (equation 2.2) but to simplify the analysis we drop the technology dimension. This leaves us with an establishment-pair fixed effect (ah) and an affiliate-country-time fixed effect ($c(a)y$). In some specifications, we also include an inventor fixed effect (i), controlling for any time-invariant inventor characteristics.

The inventor-level analysis confirms the findings of the affiliate-level analysis. The estimates in table 2.6 show that liberalization matters for cross-border collaboration when the overlap in business hours is high. Importantly, the results remain similar after including inventor fixed effects. This implies that our earlier results were not driven by a change in the composition of employees at the establishment level around liberalization episodes, but rather that individual inventors start collaborating more with HQ after liberalization.

Table 2.6: Telecom Liberalization and Collaboration - Inventor Level

	Share Collaboration			
	(1)	(2)	(3)	(4)
Liberalization	0.0030	-0.0202		
	(0.0195)	(0.0147)		
× Overlap	0.0242***	0.0220***		
	(0.0041)	(0.0034)		
× 0-2 hours			0.0182	-0.0101
			(0.0217)	(0.0156)
× 2-4 hours			0.0340	0.0205
			(0.0298)	(0.0326)
× 4-6 hours			0.0255	0.0007
			(0.0254)	(0.0246)
× 6-8 hours			0.0514**	0.0338**
			(0.0199)	(0.0165)
Observations	526,808	526,808	526,808	526,808
R^2	0.54	0.77	0.54	0.77
Dependent var. mean	0.16	0.16	0.16	0.16
Fixed Effects				
Establishment Pair	✓	✓	✓	✓
Year-Host Country	✓	✓	✓	✓
Inventor		✓		✓

Notes: This table reports the results of estimating equation 2.6. Standard errors are clustered at the country pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

6.2. *Star Inventors*

In the final part of our study, we explore heterogeneous impacts on inventors with different patenting histories. As discussed in the introduction, a priori the effects on star inventors versus their less successful peers are ambiguous. On the one hand, inventors with lesser track records may stand to benefit more from international collaborations and thus may be willing to incur the high costs of international communication and the inconvenience of time zone differential. On the other hand, the presence of such costs, which must be

partially born by their collaborators located in HQ, makes them less attractive as team members.

To examine these differential impacts we estimate the following equation:

$$
\begin{aligned}
\text{Collaboration}_{iahy} =& \beta_1 \times \text{Liberalization}_{c(a)c(h)y} + \\
&+ \beta_2 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Overlap}_{ah} + \\
&+ \beta_3 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Star}_i + \\
&+ \beta_4 \times \text{Liberalization}_{c(a)c(h)y} \times \text{Overlap}_{ah} \times \text{Star}_i + \\
&+ \beta_5 \times \text{Star}_i + \\
&+ \beta_6 \times \text{Overlap}_{ah} \times \text{Star}_i + \text{FE}_{ah} + \text{FE}_{c(a)y} + \epsilon_{iahy}
\end{aligned} \tag{2.7}
$$

This equation is akin to the previous equation , the only difference being that we introduce an inventor-level indicator for being highly productive (Star_i) and all interaction terms that are not captured by fixed effects. Note that we estimate this specification with a reduced sample, since our definition of Star_i requires an inventor to be present in the 1980s. The coefficient of interest, β_4, captures whether the increase in collaboration for inventors in high-overlap affiliates differs according to their initial track record.

We find a negative estimate for β_4 which is significant at the 5 percent level. This is independent of the way we define star inventors and is robust to including inventor fixed effects, controlling for any time-invariant differences between inventors. This means that less experienced inventors are more sensitive to an exogenous change in communication frictions. We also find some evidence that star inventors a priori collaborate more ($\hat{\beta}_5 > 0$).

Table 2.7: Telecom Liberalization and Collaboration - Inventor Heteregoneity

	Share Collaboration					
Star Inventor Definition		Top 25%		Top 10%	Continuous	Top 25% Cited
	(1)	(2)	(3)	(4)	(5)	(6)
Liberalization	-0.0133	-0.0139	-0.0233	-0.0129	-0.0151	-0.0132
	(0.0127)	(0.0125)	(0.0150)	(0.0129)	(0.0128)	(0.0128)
× Overlap	0.0168***	0.0189***	0.0204***	0.0179***	0.0205***	0.0184***
	(0.0044)	(0.0050)	(0.0037)	(0.0042)	(0.0051)	(0.0047)
× Star Inventor		0.0026	0.0104	-0.0024	0.0015	0.0005
		(0.0052)	(0.0104)	(0.0042)	(0.0018)	(0.0042)
× Overlap × Star Inventor		-0.0096**	-0.0146**	-0.0125**	-0.0033**	-0.0073**
		(0.0042)	(0.0064)	(0.0051)	(0.0015)	(0.0035)
Star Inventor		0.0036*		0.0076***	0.0009	0.0050*
		(0.0019)		(0.0018)	(0.0012)	(0.0026)
× Overlap		0.0010	0.0030	0.0018	0.0008	0.0003
		(0.0022)	(0.0054)	(0.0038)	(0.0010)	(0.0027)
Observations	149,609	149,609	149,609	149,609	149,609	149,609
R^2	0.58	0.58	0.78	0.58	0.58	0.58
Dependent var. mean	0.13	0.13	0.13	0.13	0.13	0.13
Fixed Effects						
Establishment Pair	✓	✓	✓	✓	✓	✓
Year-Host Country	✓	✓	✓	✓	✓	✓
Inventor			✓			

Notes: This table reports the results of estimating equation 2.7. Standard errors are clustered at the country pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

7. Conclusion

Using a newly constructed dataset, this paper studies innovation inside multinational firms, which frequently involves inventors based outside the HQ country. We analyze communication frictions that affect cross-border inventor collaboration and thereby the diffusion of knowledge inside the firm.

We first show that business hour overlap is crucial for collaboration between affiliate and HQ-based inventors. Inventors in technology classes involving experimentation, potentially requiring a higher frequency of communication, are most sensitive to this communication cost.

In addition to this cross-sectional observation, we study the evolution of collaboration across time. We focus on the liberalization of the telecom sector during the 1980s and 1990s. The price of international phone calls decreased drastically as a consequence, leading to an increase of cross-border communication. We find that affiliate-based inventors start collaborating with HQ alongside, especially in experimental technologies at affiliates which are temporally close to HQ. This finding shows that while cheap communication matters, it cannot overcome underlying geographic features. Recent technological advances such as video calls using Zoom and other platforms are likely subject to the same limitations.

Finally, we show that reducing frictions to cross-border knowledge creation inside multinationals is less important for star inventors than for others. This may indicate that reducing communication frictions benefits a broad set of less experienced inventors in FDI host countries, potentially reducing the knowledge gap between them and star inventors.

References

Akcigit, U., Caicedo, S., Miguelez, E., Stantcheva, S., and Sterzi, V. (2018). Dancing with the stars: Innovation through interactions. Technical report, National Bureau of Economic Research.

Antràs, P., Garicano, L., and Rossi-Hansberg, E. (2006). Offshoring in a knowledge economy. *The Quarterly Journal of Economics*, 121(1):31–77.

Bahar, D. (2020). The hardships of long distance relationships: time zone proximity and the location of mnc's knowledge-intensive activities. *Journal of International Economics*, 125:103311.

Bilir, L. K. and Morales, E. (2020). Innovation in the global firm. *Journal of Political Economy*, 128(4):1566–1625.

Blonigen, B. A., Cristea, A., and Lee, D. (2020). Evidence for the effect of monitoring costs on foreign direct investment. *Journal of Economic Behavior & Organization*, 177:601–617.

Bloom, N., Jones, C. I., Van Reenen, J., and Webb, M. (2020). Are ideas getting harder to find? *American Economic Review*, 110(4):1104–44.

Bøler, E. A., Javorcik, B., and Ulltveit-Moe, K. H. (2018). Working across time zones: Exporters and the gender wage gap. *Journal of International Economics*, 111:122–133.

Borusyak, K., Jaravel, X., and Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Boylaud, O. and Nicoletti, G. (2000). Regulation, market structure and performance in telecommunications.

Catalini, C., Fons-Rosen, C., and Gaulé, P. (2020). How do travel costs shape collaboration? *Management Science*, 66(8):3340–3360.

Chauvin, J., Choudhury, P., and Fang, T. P. (2020). The effects of temporal distance on intra-firm communication: Evidence from daylight savings time. *Harvard Business School Technology & Operations Mgt. Unit Working Paper*, (21-052).

Dachs, B., Kampik, F., Scherngell, T., Zahradnik, G., Hanzl-Weiss, D., Hunya, G., Foster, N., Leitner, S., Stehrer, R., and Urban, W. (2012). Internationalisation of business investments in R&D and analysis of their economic impact. *Luxembourg: European Commission*.

Foley, C. F. and Kerr, W. R. (2013). Ethnic innovation and us multinational firm activity. *Management Science*, 59(7):1529–1544.

Ganglmair, B., Robinson, W. K., and Seeligson, M. (2022). The rise of process claims: Evidence from a century of us patents. *ZEW-Centre for European Economic Research Discussion Paper*, (22-011).

Griffith, R., Miller, H., and O'Connell, M. (2014). Ownership of intellectual property and corporate taxation. *Journal of Public Economics*, 112:12–23.

Gumpert, A. (2018). The organization of knowledge in multinational firms. *Journal of the European Economic Association*, 16(6):1929–1976.

Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus fdi with heterogeneous firms. *American economic review*, 94(1):300–316.

Keller, W. (2002). Geographic localization of international technology diffusion. *American Economic Review*, 92(1):120–142.

Keller, W. (2004). International technology diffusion. *Journal of Economic Literature*, 42(3):752–782.

Keller, W. and Yeaple, S. R. (2013). The gravity of knowledge. *American Economic Review*, 103(4):1414–44.

Kerr, S. P. and Kerr, W. R. (2018). Global collaborative patents. *The Economic Journal*, 128(612):235–272.

Kogut, B. and Zander, U. (1993). Knowledge of the firm and the evolutionary theory of the multinational corporation. *Journal of international business studies*, 24(4):625–645.

Kvande, E. (2009). Work–life balance for fathers in globalized knowledge work. some insights from the norwegian context. *Gender, Work & Organization*, 16(1):58–72.

Marx, M. and Fuegi, A. (2020). Reliance on science: Worldwide front-page patent citations to scientific articles. *Strategic Management Journal*, 41(9):1572–1594.

OECD (1994). International telecommunication tariffs. (6).

OECD (1999). *OECD Communications Outlook 1999*.

OECD (2003). Trends in international calling prices in oecd countries.

Oldenski, L. (2012). Export versus fdi and the communication of complex information. *Journal of International Economics*, 87(2):312–322.

Pauly, S. and Stipanicic, F. (2021). The creation and diffusion of knowledge: Evidence from the jet age. Technical report, CEPREMAP.

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2022). What's trending in difference-in-differences? a synthesis of the recent econometrics literature. *arXiv preprint arXiv:2201.01194*.

Stein, E. and Daude, C. (2007). Longitude matters: Time zones and the location of foreign direct investment. *Journal of International Economics*, 71(1):96–112.

UNCTAD (2005). World investment report.

Appendix

A. Details on Orbis IP (matching patents to firms)

We provide details on the matching of patents to firms in this appendix section.

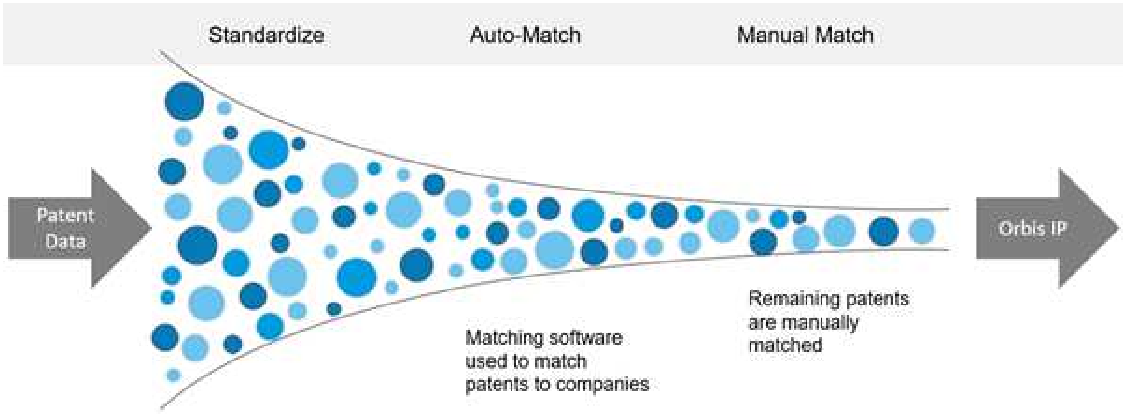
Orbis IP sources its patent information from LexisNexis. At the time of our access to Orbis IP in October-November 2020, the database had approximately 110 million patent documents, covering information from patent offices in 109 countries. Bureau van Dijk (BvD) describes their procedure of matching patents to firms as follows.

At BvD, the matching process begins with processing xml documents to extract the assignee details and company details from BvD's global company database to be matched to include the following fields:

- Company name
- Local name
- Street (not to match but to detect duplicates)
- Postal Code (not to match but to detect duplicates)
- City (important for the matching when given by the IP
- Region in country (not to match but to detect duplicates)
- Country ISO Code
- Email/Website (as a confirmation)
- Category of company
- Listed/delisted
- Status (active/inactive)
- Company ID (for duplicates)
- Source
- BvD ID
- Activity description

The above data to be matched follows the below process first through automated fuzzy logic-based matching and then through manual matching:

Fig. 2.6. Orbis IP procedure for matching patents to firms



Source: Orbis IP / Bureau van Dijk.

A.1. Automated Matching Tool (Fuzzy Logic)

BvD's fuzzy matching takes place in three steps:

1. Normalisation: transforms the original query to match the database candidates as far as possible.
2. Candidate selection: retrieves the best potential candidates from the normalised query.
3. Candidate evaluation: evaluates the distance of each candidate from the normalised query, and filters the candidates to select the best matches.

These three steps are explained in more detail below.

1. Normalisation

The goal of the normalisation process is to reduce the differences between the original query and the database entities to a minimum. When entities are added to the database, they are normalised with the same process that is used to normalise the query, ensuring minimal differences.

Different normalisation processes take place (lowercase, remove diacritics, etc). Some are specific to the type of entity: for companies, the legal form will be normalised (i.e. Limited becomes ltd) and for individuals the first name will be normalised using hypocorism.

Note that the normalisation process for one entity can lead to multiple searches: consider a single query to retrieve Thierry Henry (French football player). Thierry and Henry are both French first names and last names, so as part of the normalisation process it performs two distinct searches: a) one for first name Thierry, last name Henry b) and one for first name Henry, last name Thierry.

2. Candidate selection

The goal of the selection step is to retrieve - as quickly as possible - all potential candidates in the database. When handling a large quantity of data it isn't possible to check every entity of the database in order to retrieve only the best candidates, therefore a fuzzy algorithm is used in order to locate within the database only those candidates that "should be" the best ones, and to evaluate only those candidates.

All such algorithms come at a cost: there is a trade-off to be made between the "recall" and the "precision". Globally, the recall will measure the fact that the selected candidates always contain the true match, while the precision will measure the fact that the proposed candidate are valid matches, and will limit false positives in the results. No algorithm will guarantee a 100% recall, except if you evaluate all the candidates present in the database one by one, which isn't possible when handling a large quantity of data.

The algorithm implemented for the candidate selection is based on the N-Gram algorithm, where a query is split into different blocks of letters (or grams) of a given size (N=3). The potential candidates are then those candidates that share as much N-Gram as possible with the normalised query. This algorithm is language agnostic. For example, if you misspell "Bureau van Dijk" and query the system for "Buro van Dijk", the decomposition in N-Gram of both names (N=3) would give:

BUR URO VAN DIJ IJK

BUR URE REA EAU VAN DIJ IJK

and would retrieve BUR, VAN, DIJ, IJK in common to both queries, leading to the selection of the candidate.

Note that the major flaw of the N-Gram algorithm is working with "small" words, indeed making a mistake in a three-letter word using 3 as the gram size would never select

the correct three-letter candidate as no gram will be common (i.e. if one were to type BNW intending to retrieve BMW). Therefore, small words are handled differently and the applicable method uses an algorithm based on the edit distance.

3. Candidate evaluation

In the last step of the process, the goal is to sort and reduce the candidates to the distance threshold value specified by the user. The computed similarity is based on the edit distance.

A.2. Manual Matching Software

The BvD Matching Software application is designed to supplement the automatic matching functionality that is typically used to correlate records found in the patent databases to records in BvD's company database. Records that cannot be matched automatically are presented to BvD Matching Software users so that they can accept or reject possible matches manually. The likelihood of a match is declared for each record in the application's interface. The final purpose of the matching process is to link patent records to an appropriate BvD identification number.

BvD uses the following fields to calculate a matching score between patent and BvD company databases: Assignee name, Street, City, Postal code, Country (or ISO code). The registered/legal address of the companies is used. BvD takes into consideration the current name of the companies, alongside their previous names and 'also known as' names. There is no limit in the number of characters for a name (or any other field).

1. Normalisation To identify similarities and enable high scoring matches, the matching system uses n-gram indexes, normalisation rules, and data dictionaries. The system can handle spelling mistakes, typos, word orders, special characters, context of words as it relates to a specific field/country, etc.

Specific normalisation rules are defined for each possible matching data field, and for each country. The normalisation rules are applied on BvD records (in the Orbis database) and patent records, hence the normalised values are taken into account for comparison/matching. For example, punctuation marks are replaced by blanks, legal forms are standardized, non-relevant words are ignored, synonyms are converted to a simple form, accented characters are converted into non-accented characters.

BvD Matching Software is Unicode compliant. It supports local characters such as Chinese, Cyrillic, Hungarian, German, Polish, Arabic, etc., and for some specific languages BvD uses transliteration to transform local characters into common and comparable characters enabling cross-alphabet matching.

2. Calculation After the data have been normalised, the match score is calculated using an algorithm based on the following principle. For each field of a given record, calculate the % of accuracy of the match between the BvD and patent records by using proximity calculations. Then, take a weighted average of all the % of proximities. Weights are calculated automatically; they are not fixed. The weight of a criterion is based on the probability of finding a company in the BvD database corresponding to the criterion being searched. This means the weight of each criterion depends on the number of occurrences in the BvD database and therefore, may differ from one release to another. The more occurrences that are found, the less the field is significant.

The automated process produces a matching score for each record. A quality indicator uses the following scoring criteria:

Fig. 2.7. Orbis IP patent-to-firm matching score indicator

A	Excellent	total score >= 95 %
B	Good	total score between 85 and 94 %
C	Fair	total score between 75 and 84 %
D	Weak	total score between 60 and 74 %
E	Poor	total score < 60 %

Source: Orbis IP / Bureau van Dijk.

The higher the score, the better the data match. Candidates with a poor matching quality (E) may be irrelevant. Any match with less than a 70% score is pushed into the manual matching pipeline.

B. Sample Coverage

This section of the appendix provides details on the sample coverage of our dataset.

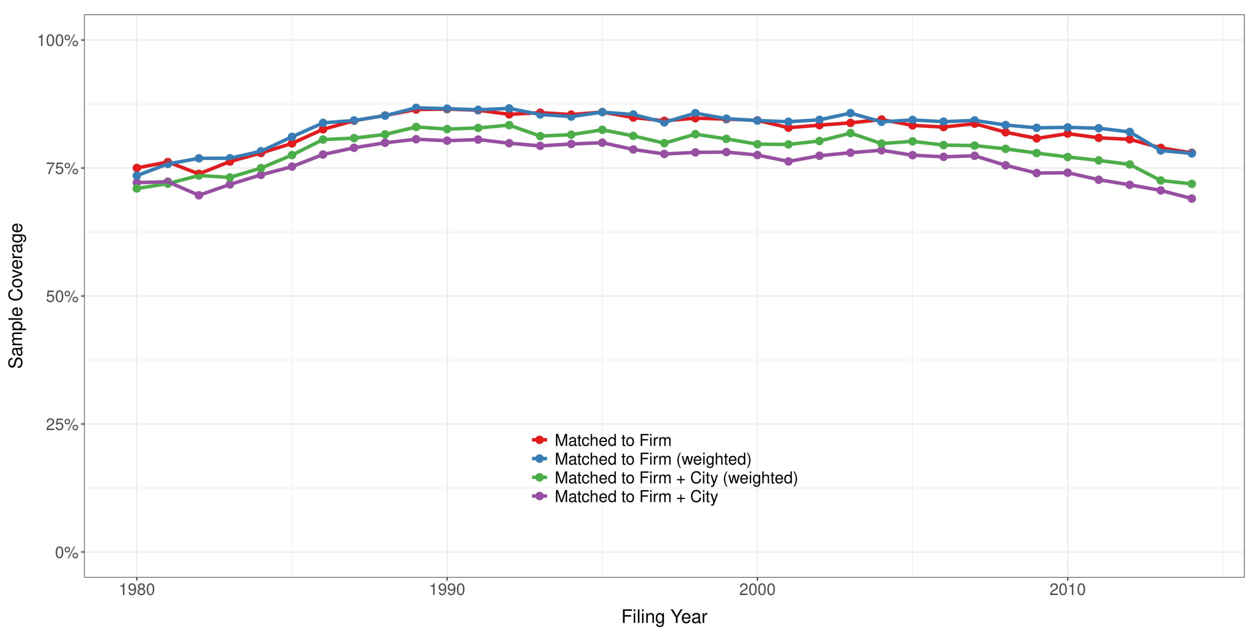
To determine the sample coverage, we first need to calculate the relevant total number of granted triadic patents. We focus on patents that are present in both PATSTAT and

PatentsView. There are some differences in what these two datasets contain. Noteably, PATSTAT includes design patents only after 2001. Since we cannot use PATSTAT to determine the members of the extended patent family for these early design patents, they are not considered in the relevant total number of patents we try to match.

The most relevant variables for our analysis are the Orbis firm identifier and the inventor location (at the city level) from PatentsView. We focus on inventor location rather than assignee location to capture where innovation takes place physically. Both the firm identifier and the inventor location can be missing.

One patent often lists multiple inventors and sometimes also multiple assignees. When we compute sample coverage at the patent level we flag a patent as covered if the patent is matched to at least one assignee and one inventor location. Based on this methodology we find a coverage of around 83% (or 84% when weighting by forward citations). The coverage is slightly lower with 79% when we look at patents that are granted in the US and filed but not necessarily granted at EPO and JPO. In figure 2.8 we plot the sample coverage for granted triadic patents by year. It remains relatively stable between 75% and 85% with the lowest coverage at the beginning and the end of the sample.

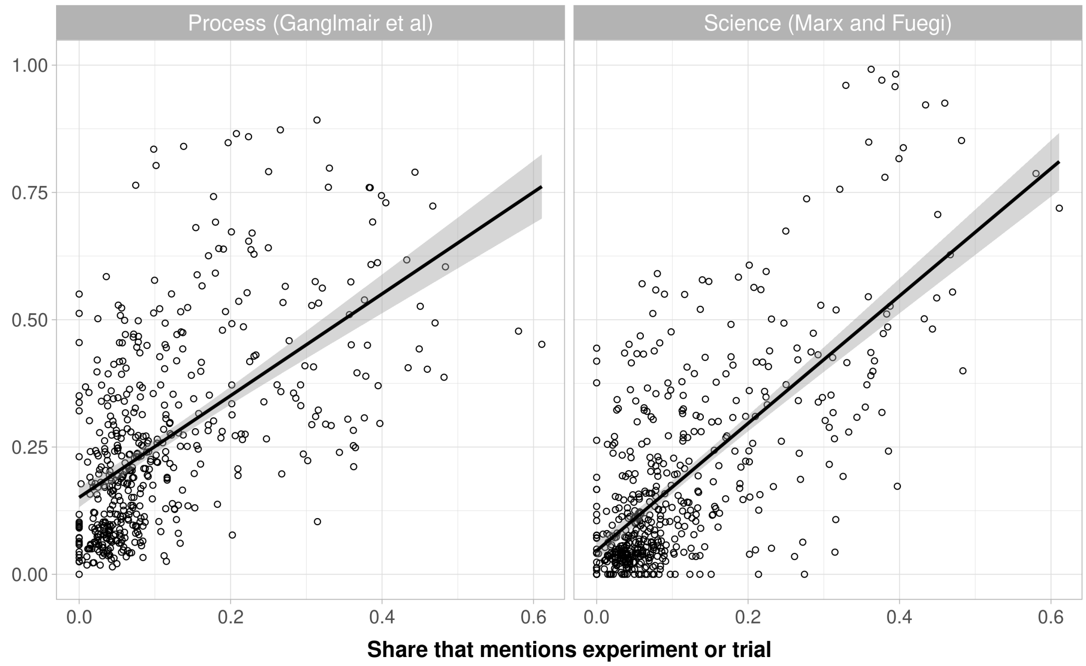
Fig. 2.8. Sample Coverage



Notes: This figure shows the yearly share of granted triadic patents that our sample covers.

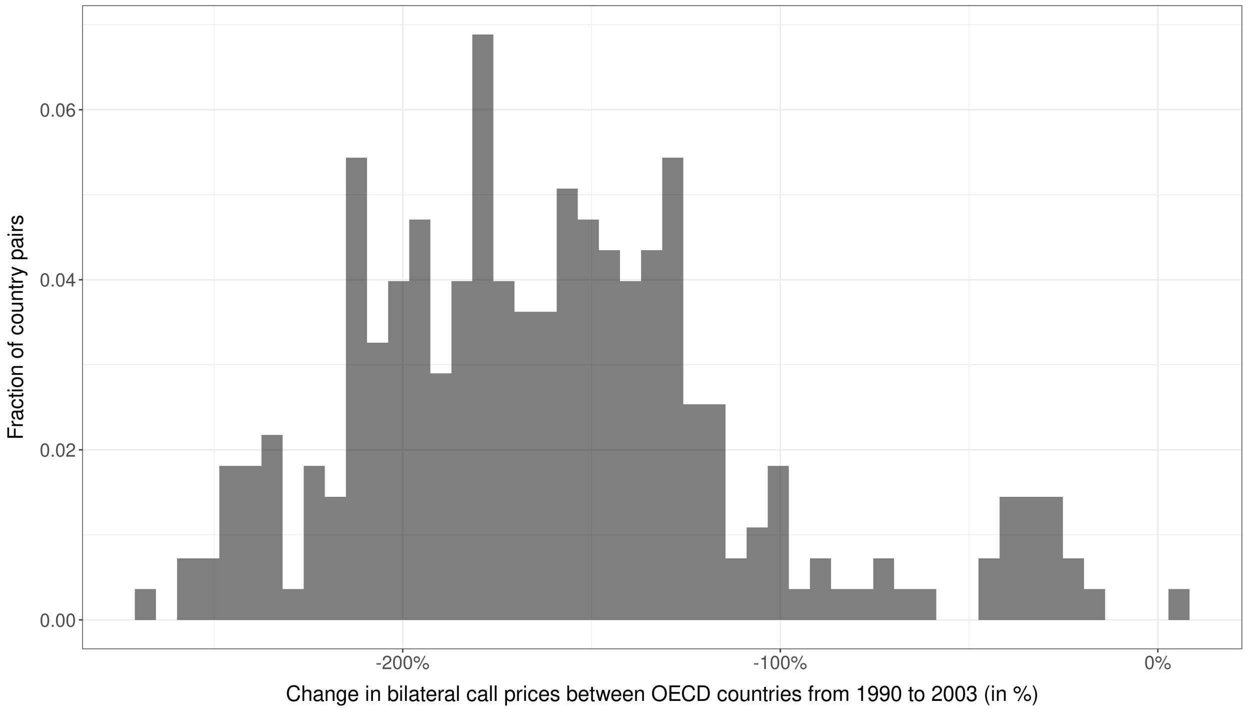
C. Additional Figures

Fig. 2.9. Experimental Technology Classes vs other Classifications



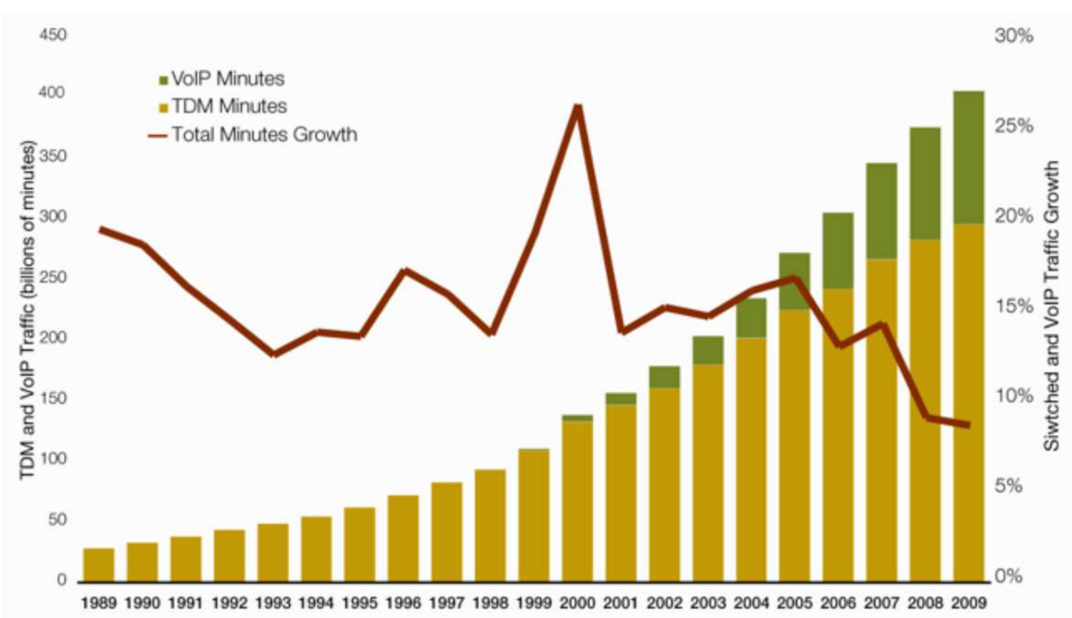
Notes: Each dot represents a technology class. The x-axis captures our measure of experimental technology class, i.e. the share mentioning experiments or trials. In panel A, the y-axis plots the share of patents containing a process claim (based on data from Ganglmair et al. (2022)) and in panel B the y-axis depicts the share of patents citing a scientific article (based on data from Marx and Fuegi (2020).)

Fig. 2.10. Rate Change



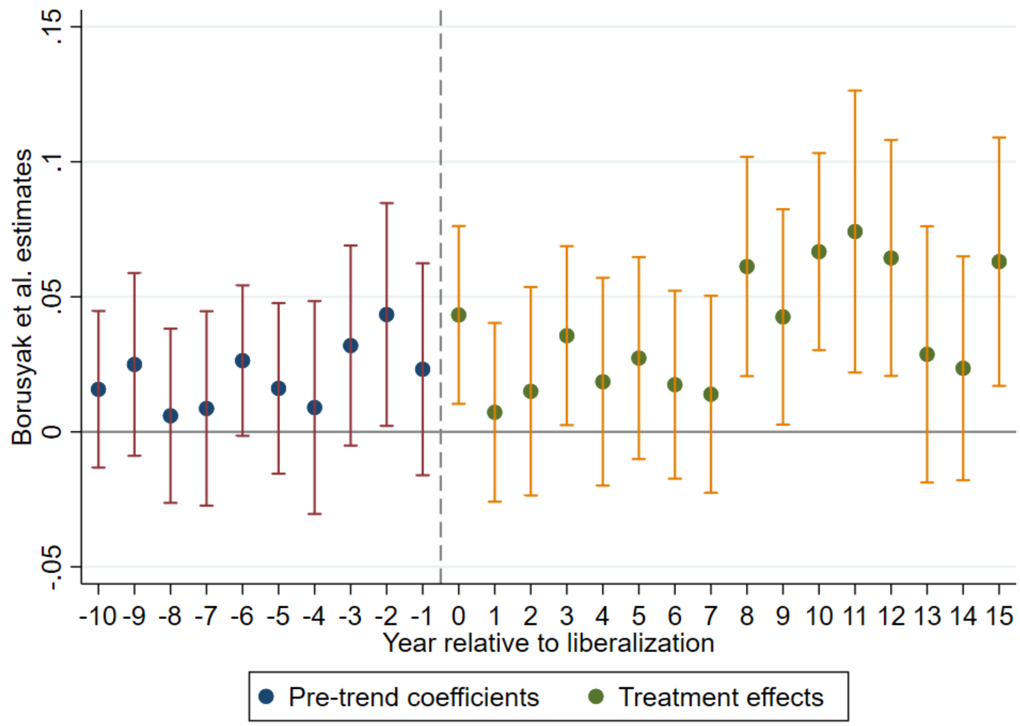
Source: OECD.

Fig. 2.11. Volume of international calls

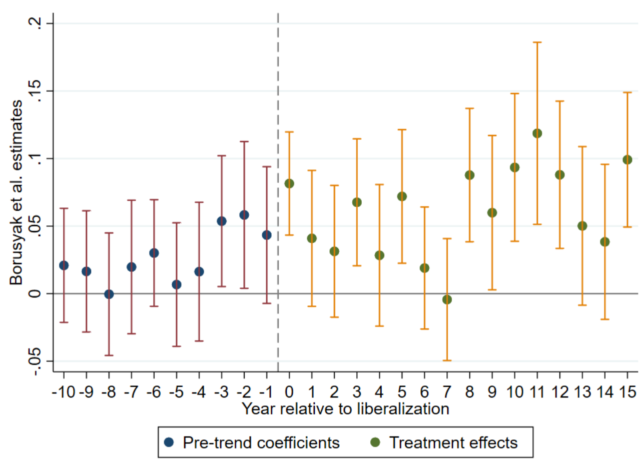


Notes: This picture is taken from a 2009 Telegeography report. It shows the rise of international call volumnes during our sample period.

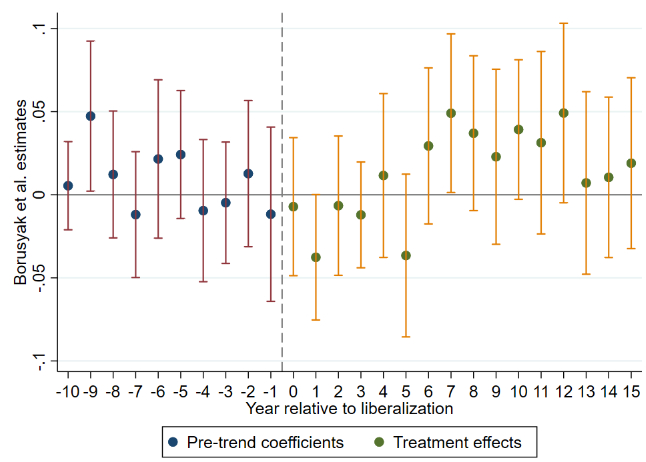
Fig. 2.12. Event Study of the Impact of Telecom Liberalization on Collaboration - Borusyak et al. (2021) estimates



(a) Full sample



(b) Experimental



(c) Non-Experimental

Notes: This figure reports the results of estimating equation 2.3, including unit and time fixed effects only. Standard errors are clustered at the country-pair level. 95% confidence intervals are shown.

D. Additional Tables

Table 2.8: Share of experiment-intensive Technologies

IPC Class	Share Experimental	Patent Count
Chemistry, Metallurgy	0.98	242455
Human Necessities	0.63	172939
Electricity	0.00	122025
Performing Operations, Transporting	0.34	120374
Physics	0.28	120271
Mechanical Engineering, Lighting, Heating, Weapons, Blasting	0.12	42462
Textiles, Paper	0.76	13287
Fixed Constructions	0.00	6159

Notes: This table shows the share of experimental (4-digit) technology classes by the more aggregate 1-digit IPC classes.

Table 2.9: Classification of Top-15 Technologies

Technology Code	Technology	Share Mentioning Experiment/Trial	Classified as Experimental	Patent Count
C07D	heterocyclic compounds	0.48	Yes	28121
C07C	acyclic or carbocyclic compounds	0.47	Yes	14001
A61P	therapeutic activity of chemical compounds or medicinal preparations	0.43	Yes	37235
C12N	micro-organisms or enzymes; compositions thereof; propagating, preserving, or maintaining micro-organisms; mutation or genetic engineering; culture media	0.39	Yes	18183
C08G	macromolecular compounds obtained otherwise than by reactions only involving carbon-to-carbon unsaturated bonds	0.37	Yes	11627
A61K	preparations for medical, dental, or toilet purposes	0.36	Yes	49856
C08L	compositions of macromolecular compounds	0.36	Yes	16546
C07K	peptides	0.36	Yes	15852
G01N	investigating or analysing materials by determining their chemical or physical properties	0.22	Yes	21043
A61B	diagnosis; surgery; identification	0.11	No	15806
A61F	filters implantable into blood vessels; prostheses; devices providing patency to, or preventing collapsing of, tubular structures of the body, e.g. stents; orthopaedic, nursing or contraceptive devices; fomentation; treatment or protection of eyes or ears; bandages, dressings or absorbent pads; first-aid kits	0.09	No	10962
H04L	transmission of digital information, e.g. telegraphic communication	0.06	No	17932
G06F	electric digital data processing	0.05	No	17506
H01L	semiconductor devices; electric solid state devices not otherwise provided for	0.05	No	13193

Notes: This table shows our classification of experimental technology classes for the 15 most frequent.

Table 2.10: Liberalization and Call Prices

HQ Country	Number of patent families	Only HQ inventors (in %)	Only foreign affiliate inventors (in %)	HQ-affiliate collaboration (in %)
JP	110277	88.71	5.82	5.47
US	76772	60.60	16.33	23.06
DE	34077	58.63	12.92	28.45
FR	17353	50.81	24.91	24.28
KR	13555	77.29	6.54	16.17
GB	7031	34.79	32.20	33.01
SE	6384	48.29	26.10	25.61
IT	5311	55.71	13.22	31.07
NL	4796	8.42	48.81	42.76
CH	4571	5.73	55.83	38.44

Table 2.11: Collaboration increases Patent Quality

	log(1+Citations)		Cites Scientific Paper	
	(1)	(2)	(3)	(4)
Collaboration with HQ country	1.189***	0.1008***	0.1939***	0.0330***
	(0.1307)	(0.0216)	(0.0084)	(0.0072)
log(Inventors)		0.6491***		0.0811***
		(0.0341)		(0.0106)
Observations	1,314,466	1,314,466	1,314,466	1,314,466
R^2	0.32	0.95	0.07	0.84
Dependent variable mean	3.7	3.7	0.69	0.69
Fixed Effects				
Filing Year	✓	✓	✓	✓
Inventor		✓		✓
GUO		✓		✓
Country-Technology		✓		✓

Notes: This table regresses the number of forward citations ($\log(1+x)$-transformed) on a binary variable indicating whether a patent is filed in collaboration with HQ inventors, controlling for a variety of fixed effects. In columns 2 and 4 we also control for the number of inventors (log-transformed).

Table 2.12: Collaboration and Business Hour Overlap

	Share Collaboration			
	(1)	(2)	(3)	(4)
Overlap	0.0646***	0.0593***	0.0557***	0.0584***
	(0.0019)	(0.0020)	(0.0023)	(0.0021)
× Experimental		0.0111***		0.0117***
		(0.0028)		(0.0032)
× Share Experimental			0.0446***	
			(0.0098)	
× Cites Science				-0.0042
				(0.0034)
× Process				0.0053*
				(0.0029)
Observations	316,875	316,875	316,875	316,875
R^2	0.47	0.47	0.47	0.47
Dependent var. mean	0.36	0.36	0.36	0.36
Fixed Effects				
GUO	✓	✓	✓	✓
Host Country-Technology	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.1. Standard errors are clustered at the location pair-technology level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 2.13: Year of Liberalization of International Calls

Year of Liberalization	Country
1984	United States
1986	United Kingdom
1987	Japan
1990	New Zealand
1991	Australia
1992	Canada, Sweden
1993	Finland
1996	Denmark, Korea, Mexico
1997	Netherlands
1998	Austria, Belgium, France, Germany, Ireland
	Italy, Luxembourg, Norway, Spain, Switzerland
2000	Czech Republic, Portugal
2001	Greece
2002	Hungary
2006	Turkey

Notes: This table shows the year of liberalization of the telecommunication sector, as reported in Boylaud and Nicoletti (2000).

Table 2.14: Liberalization - 1980s Establishments

	Share Collaboration					
	Full Sample	Experimental	Non-experim.	Full Sample	Experimental	Non-experim.
	(1)	(2)	(3)	(4)	(5)	(6)
Liberalization	0.0112	0.0166	0.0046			
	(0.0167)	(0.0191)	(0.0220)			
× Overlap	0.0245***	0.0316***	0.0118**			
	(0.0069)	(0.0092)	(0.0056)			
× 0-2 hours				0.0261	0.0367**	0.0094
				(0.0158)	(0.0168)	(0.0218)
× 2-4 hours				0.0417	0.0359	0.0562**
				(0.0340)	(0.0468)	(0.0221)
× 4-6 hours				0.0312	0.0243	0.0455*
				(0.0199)	(0.0218)	(0.0232)
× 6-8 hours				0.0623***	0.0888***	0.0180
				(0.0156)	(0.0187)	(0.0190)
Observations	316,967	170,235	146,732	316,967	170,235	146,732
R^2	0.74	0.69	0.80	0.74	0.69	0.80
Dependent var. mean	0.19	0.20	0.16	0.19	0.20	0.16
Fixed Effects						
Est. Pair-Technology	✓	✓	✓	✓	✓	✓
Year-Host Country	✓	✓	✓	✓	✓	✓
Year-Technology	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.2, keeping only affiliates that already existed in the 1980s. Standard errors are clustered at the country pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 2.15: Liberalization - Source Country-Time Fixed Effect

	Share Collaboration					
	Full Sample	Experimental	Non-experim.	Full Sample	Experimental	Non-experim.
	(1)	(2)	(3)	(4)	(5)	(6)
Liberalization	-0.0072	-0.0081	-0.0049			
	(0.0072)	(0.0081)	(0.0121)			
$\times$ Overlap	0.0176***	0.0263***	0.0004			
	(0.0055)	(0.0071)	(0.0057)			
$\times$ 0-2 hours				-0.0015	0.0013	-0.0064
				(0.0064)	(0.0070)	(0.0113)
$\times$ 2-4 hours				0.0261	0.0356	0.0135
				(0.0308)	(0.0406)	(0.0167)
$\times$ 4-6 hours				0.0217	0.0294	0.0083
				(0.0180)	(0.0211)	(0.0198)
$\times$ 6-8 hours				0.0380***	0.0560***	0.0007
				(0.0097)	(0.0125)	(0.0147)
Observations	575,780	293,241	282,539	575,780	293,241	282,539
R^2	0.84	0.81	0.87	0.84	0.81	0.87
Dependent var. mean	0.27	0.28	0.26	0.27	0.28	0.26
Fixed Effects						
Est. Pair-Technology	✓	✓	✓	✓	✓	✓
Year-Source Country	✓	✓	✓	✓	✓	✓
Year-Technology	✓	✓	✓	✓	✓	✓

Notes: This table reports the results of estimating equation 2.2. Standard errors are clustered at the country pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 2.16: Liberalization - Triple Interactions

	Share Collaboration	
	(1)	(2)
Liberalization	0.0099	
	(0.0175)	
× Experimental	0.0083	0.0078
	(0.0142)	(0.0156)
× Overlap	0.0103*	
	(0.0058)	
× Experimental × Overlap	0.0167**	0.0233***
	(0.0072)	(0.0070)
Observations	575,780	575,780
R^2	0.84	0.86
Dependent variable mean	0.27	0.27
Fixed Effects		
Establishment Pair-Technology	✓	✓
Year-Host Country	✓	
Year-Technology	✓	✓
Year-Location Pair		✓

Notes: This table reports the results of estimating equation 2.2. Standard errors are clustered at the country pair level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Chapter 3

Traders in Trouble on the Road to Mandalay

Abstract

This paper studies the effect of uncertainty in trade costs on exporting. By combining data on armed conflict with trade transaction data from Myanmar, I show that conflicts close to highways have a significant impact on trade flows. Firm-level export transactions via a route becoming risky because of conflict drop by around 50%. In addition, prices for export goods in Yangon's retail markets, itself far away from conflict areas, drop when conflicts put foreign market access in jeopardy by around 5%. This pass-through to local prices suggests that traders have few export alternatives and sell domestically instead. In addition, I observe that exporters of non-perishable goods lump transactions together, reducing the negative impact of conflict. Among exporters of perishable and thus riskier products, large and politically connected firms are the most resilient.

JEL codes: D22, D74, D80, F14, O12
Keywords: Trade, Uncertainty, Conflict, Development

1. Introduction

Recent events such as the Covid-19 pandemic or the Russian invasion of Ukraine have put a spotlight on the functioning of global trade networks. Sudden rises of infections in China's port cities have repeatedly disrupted container shipping, armed conflict around Ukraine's ports has made it difficult to ship crucial wheat exports to its customers. Especially in developing countries disruptions of trade routes and delays are not new phenomena. The lack of modern infrastructure makes travel time and trade costs uncertain. Unexpected landslides or heavy rain can cause road closures, insecurity can make it dangerous to travel parts of the territory and make market access unreliable.

Disruptions of trade routes can cause large losses, especially if the delivery of goods is time-sensitive as is the case for perishable goods. Insurance markets are often absent, implying that traders themselves bear the burden. How are firms affected by an increase in the risk of disruption? Are they sufficiently diversified so that they start selling to other markets? Are some firms better equipped to handle risk?

This paper studies a setting where market access is frequently disrupted. In particular, I look at exports leaving Myanmar via its land borders. Exports via Burmese land borders accounted for around 32% of total export revenue in 2014.[1] The value of exports in 2014 leaving Myanmar via Muse alone, which is the most important border crossing with China, was around one billion US dollars (or around 1.7% of GDP).

At the same time Myanmar is faced with recurring ethnic conflicts. These occur mainly in the border regions with China, Thailand and Bangladesh. This is precisely where the shipments going to the border need to pass through. To illustrate this point, the map in figure 3.1 shows the fastest route from central Yangon, the commercial capital and home to many exporters, to all border stations. In addition, the figure shows where armed conflicts took place in the 2010s. It becomes apparent that many conflicts happen close to the important trade corridors. Trucks can get stuck due to road closures or drivers refuse to pass the conflict-ridden areas. These delays can result in losses for exporters.

Relying on detailed customs data on exports and geocoded data on conflict, I show that firms decrease the number of monthly export transactions via a border station by around 50% when conflicts occur close to roads that lead there. This decline in the frequency of

[1]This figure is somewhat higher when including oil and gas which are transported via pipelines that also cross the land borders.

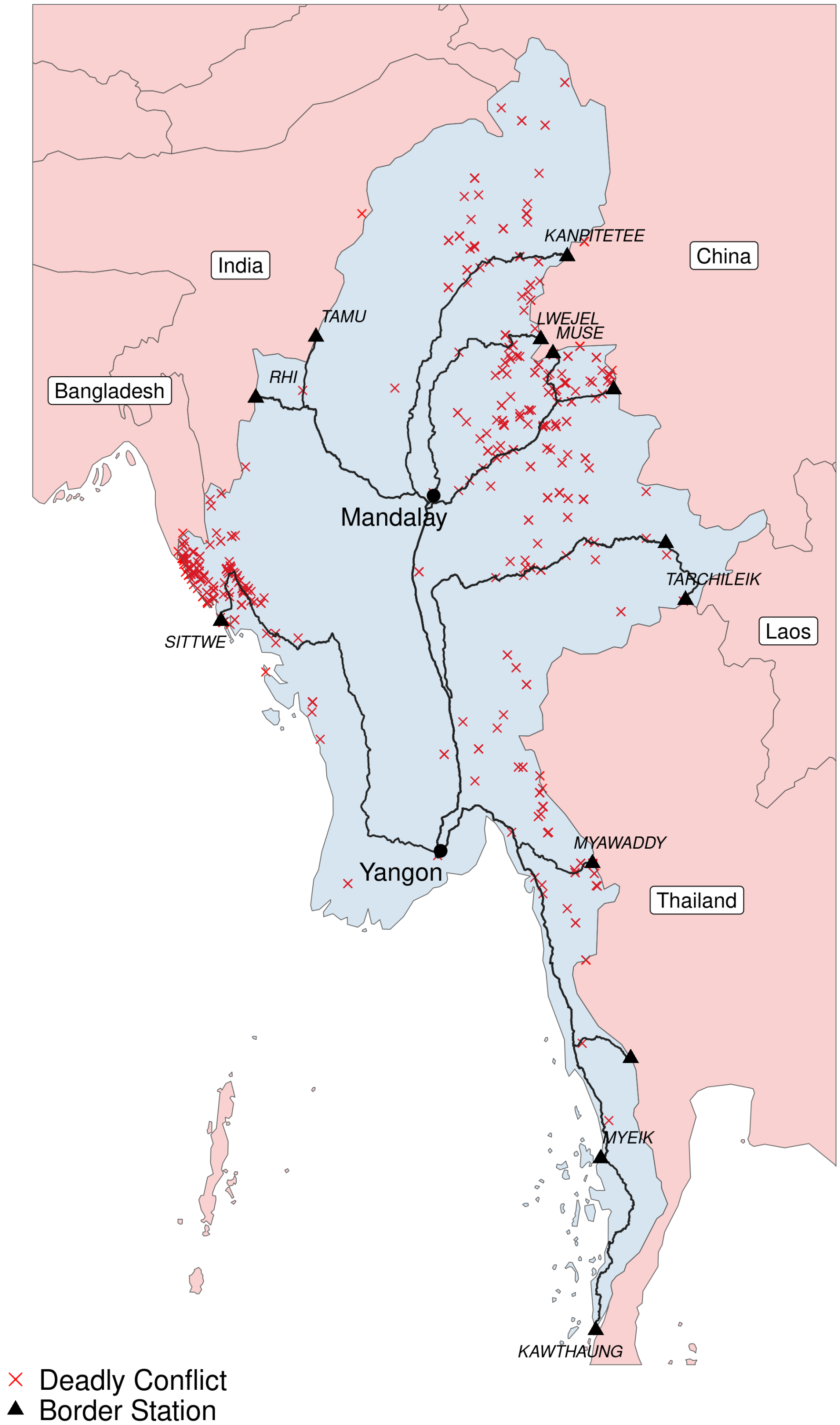


Fig. 3.1. Routes from Yangon to Myanmar's border stations and Conflict
Notes: The figure shows the fastest routes from central Yangon to each of Myanmar's border stations. I also plot the conflicts that occurred between 2010 and 2020. See section 3 for a detailed description of the data.

trade due to insecurity is accompanied by a drop in local retail prices for export goods of around 5% in Yangon, which is itself not affected by conflicts. Hence, these disruptions have an impact far beyond where they take place and lead exporters to sell their goods on domestic markets, depressing prices for these goods and leading to substantial economic losses.

I then show that instead of selling on the domestic market, some exporters seem to stockpile and lump together transactions. In particular, while all exporters reduce the number of shipments during months of turmoil, the average shipment size of non-perishable products increases. This implies that firms delay shipments and compensate partly for disruptions to market access. However, many exports are perishable and cannot easily delay transactions. I use this differential impact of conflict depending on the perishability of products to identify firm characteristics that make firms more resilient against risk.

First, I show that large firms are more likely to continue exporting in spite of conflict, with the differential being larger in riskier sectors. Thus, large firms are better able to handle the risk of losing shipments. This skews the exporter size distribution towards larger firms during conflict periods.

Second, I exploit information on political connections. Firm registry data matched with sanction lists allows identifying firms connected to the military which is a party to most conflicts in Myanmar. I show that firms which sell risky products and are connected to the military keep exporting during conflict. These firms are thus shielded from the increased risk of doing business during turmoil.

Related Literature. This paper relates to several strands of the literature. First, this study contributes to the nascent literature on how the economic impact of conflict spreads along production networks. Korovkin and Makarin (2021) show how the Russian invasion of the Donbas region in Ukraine led to the sudden removal of some firms from Ukranian supply chains. Using railway transaction data, they show how this shock propagated along the network. Similarly, Ksoll et al. (2021) study the impact of electoral violence in Kenya affected the floriculture industry. They focus on direct disruptions to supply caused by worker absence. Finally, Couttenier et al. (2022) develop a structural model to study how shocks to firms in conflict-affected areas propagate along the production network. While this literature exclusively focuses on direct disruptions to supply and how these propagate along supply chains, this paper is to the best of my knowledge the first to show that conflict has an additional effect by making trade routes insecure to pass and thereby

disrupting market access.

More generally, there is a growing literature on supply chain disruptions and how they propagate along production networks (see e.g. Barrot and Sauvagnat (2016), Boehm et al. (2019) or Lafrogne-Joussier et al. (2021) for a recent contribution). These papers typically use shocks that directly affect supply, e.g. by destroying physical capital. In addition, these papers typically rely on evidence from industrialized economies. In contrast, this paper studies disruptions in a developing country setting to trade networks that affect the links (transport routes) rather than the nodes (firms).

Second, this paper relates to a literature that studies trade costs in developing countries. Atkin and Donaldson (2015) show that domestic trade costs in Ethiopia and Nigeria is four to five times larger than in the US. I contribute to this literature by highlighting conflict as one barrier to trade in a fragile country like Myanmar. Similarly, Chiovelli et al. (2018) study the effect of land mines on economic development. Their analysis shows that removing landmines located close to transport routes in Mozambique increased economic growth in regions that depend on these routes for access to markets. Their paper studies how in a post-conflict environment landmines can make roads unusable for a sustained period of time, while I focus on how short-run disruptions affect trade and access to foreign markets while the conflict is ongoing.

Third, this paper contributes to the study of uncertainty in trade. In particular, it relates to the literature on trade policy uncertainty (TPU).[2] Most prominently, Handley and Limão (2017) study how China's WTO accession resulted in a reduction in uncertainty about future tariff hikes. This lead to an increase in Chinese exports to the US. While they analyze a reduction of uncertainty of long-run tariff policies, I study short-run disruptions of trade routes in a developing country context.

Structure. The structure of the paper is as follows. Section 2 gives a brief overview of Myanmar's ethnic conflicts and the structure of its exports. Section 3 introduces the data. In section 4 I show the direct impact of conflict on trade flows. Section 5 uses the perishability of products to identify risky products and studies the differential reaction by firms to this risk. Section 6 concludes.

[2] See Handley and Limão (2022) for a comprehensive overview of this literature.

2. Context

This section first describes the nature of Myanmar's land border exports and then gives a brief overview of ethnic conflicts that can disrupt these trading activities.

2.1. Myanmar's Exports

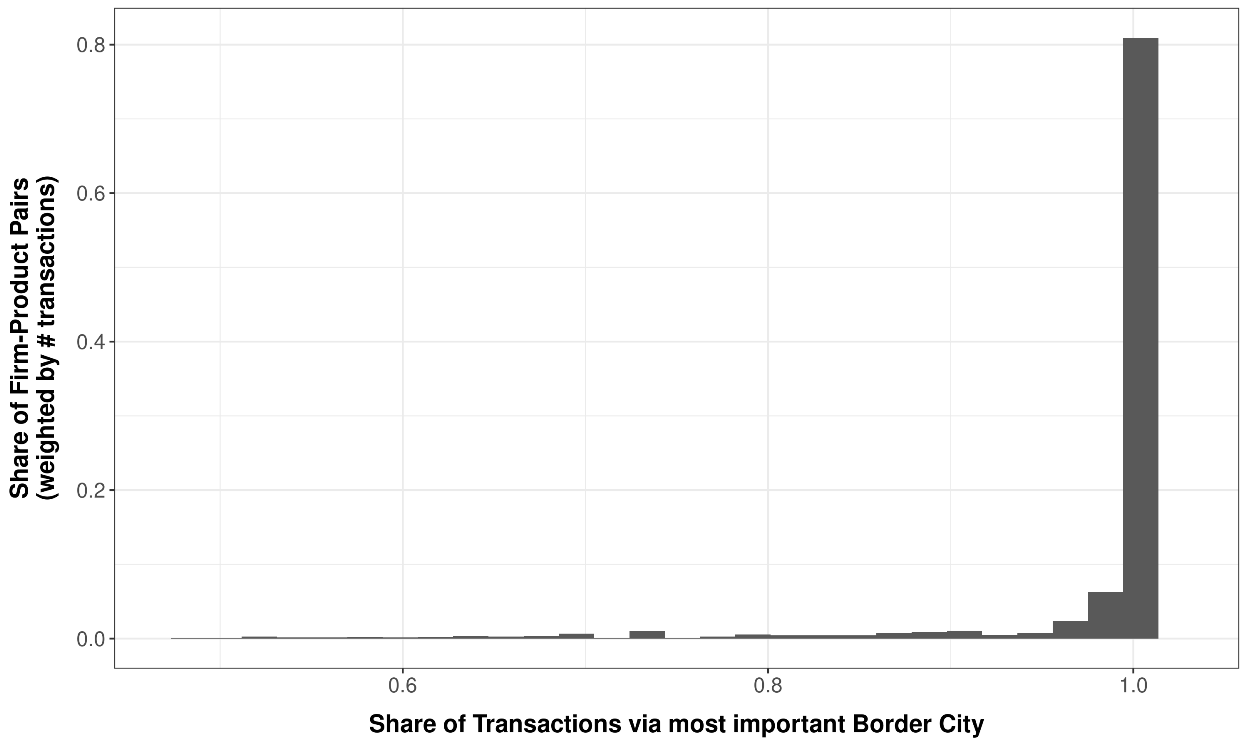
To study exporting under risk, the analysis will focus on exports leaving Myanmar via its land borders. To give an example which illustrates the context, a typical seafood exporter buys produce in Yangon, then processes the goods and ships to a border city (see the map in figure 3.1 for the location of border cities). Buyers cross the border from the neighboring country to inspect the products and agreements are usually made on the spot.[3]

An important feature of this setting is that exporters typically transact via just one border station. Figure 3.2 shows the share of export transactions via an exporter's top border city over the full time interval of the data from 2011 to 2017 for a particular product. More than 80% of firms export a product via just one border city. Note that the share of product-firm pairs is weighted by the number of transactions. The stylized fact is thus not driven by a lot of small exporters.[4]

[3]The exact export procedure varies by product, but most exports follow a similar trajectory where the produce first needs to get to a border city via Mandalay or Yangon, and is then sold to foreign buyers, often in auctions.

[4]This suggests that there are entry costs associates with border stations. See, for example, `https://www.mmtimes.com/business/23878-muse-offensive-closes-border-trading.html`. U Khin Maung Lwin of the ministry of commerce notes that moving to other export routes during conflicts is made difficult by a lack of familiar trade connections.

Fig. 3.2. Lack of Diversification



Notes: This figures plots the share of export transactions for each product product leaving Myanmar via an exporter's most important border city.

In addition to this lack of diversification, shipments are not insured. If the shipment loses in value on the way, e.g. because delays cause part of the products to rot, the exporter faces large losses of revenue. A localized disruption of the transport network can thus have large implications for firms and the economy as a whole.

Indeed land border exports are important for the Burmese economy. Overall, land border exports accounted for 35% of export revenue in 2014 (excluding oil and gas, since it is transported via pipelines). They thus account for a substantial fraction of overall exports. These consist mainly of primary commodities. Focusing on exports leaving Myanmar via its land borders, table 3.8 shows the top 10 export products in 2014, ranked by the number of transactions (again dropping oil and gas). Seafood and agricultural products are clearly dominant. Perishable goods are thus an important part of Myanmar's exports and are at the same time especially vulnerable to disruptions.

Table 3.1 shows the number of transactions and the value of products passing through the different border stations. Trade with China is heavily concentrated in Muse which is

a large border town in Northern Shan State that largely lives off the cross-border trade. It is connected to the main commercial center, Yangon, via the Mandalay-Muse highway. This highway section is crucial for Myanmar's exporters and at the same time frequently subject to disruption. Trade with Thailand is less concentrated across border stations and exporters mainly trade via three different border stations: Kawthaung, Myawaddy and Myeik. Trade with Bangladesh and India only accounts for a small fraction of Myanmar's land-border exports.

Table 3.1: Myanmar's Border Stations in 2014

Destination Country	Border Station	Transactions	Value (mln USD)
Bangladesh	SITTWE	1006	7.29
Bangladesh	MAUNG DAW	363	7.85
China	MUSE	22710	1066.54
China	CHIN SHWEHAW	1396	247.13
China	LWEJEL	345	59.34
China	KYAING TONG	186	4.76
China	KANPITETEE	123	8.72
India	TAMU	1828	26.13
India	RHI	508	9.82
Thailand	KAWTHAUNG	6646	35.94
Thailand	MYAWADDY	4618	30.48
Thailand	MYEIK	2140	91.06
Thailand	TARCHILEIK	371	8.64
Thailand	NABULAE/HTEE KHEE	38	0.32

Myanmar's exporters are concentrated in the main business centers. Around 50% of export transactions originate in Yangon, followed by Mandalay with 16%. Both of these cities are located in the central, more peaceful part of the country. Most firms are thus themselves located in an area that is usually not directly affected by conflicts but as I will show below suffer from frequent disruptions of trade routes and thus loss of market access. The map in figure 3.5 shows the spatial distribution of exporters across Myanmar.

Table 3.2: Myanmar's Land Exporters in 2014

Variable	N	Minimum	Median	Mean	Max	SD
# Transaction	459	1.00	12.00	92.11	2199.00	233.25
# HS6 Products	459	1.00	2.00	5.30	44.00	7.61
# Border Stations	459	1.00	1.00	1.43	6.00	0.90
# Destination Countries	459	1.00	1.00	1.24	3.00	0.52
Total Export Revenue (in thsd USD)	459	0.88	400.01	3494.59	64695.47	8578.38

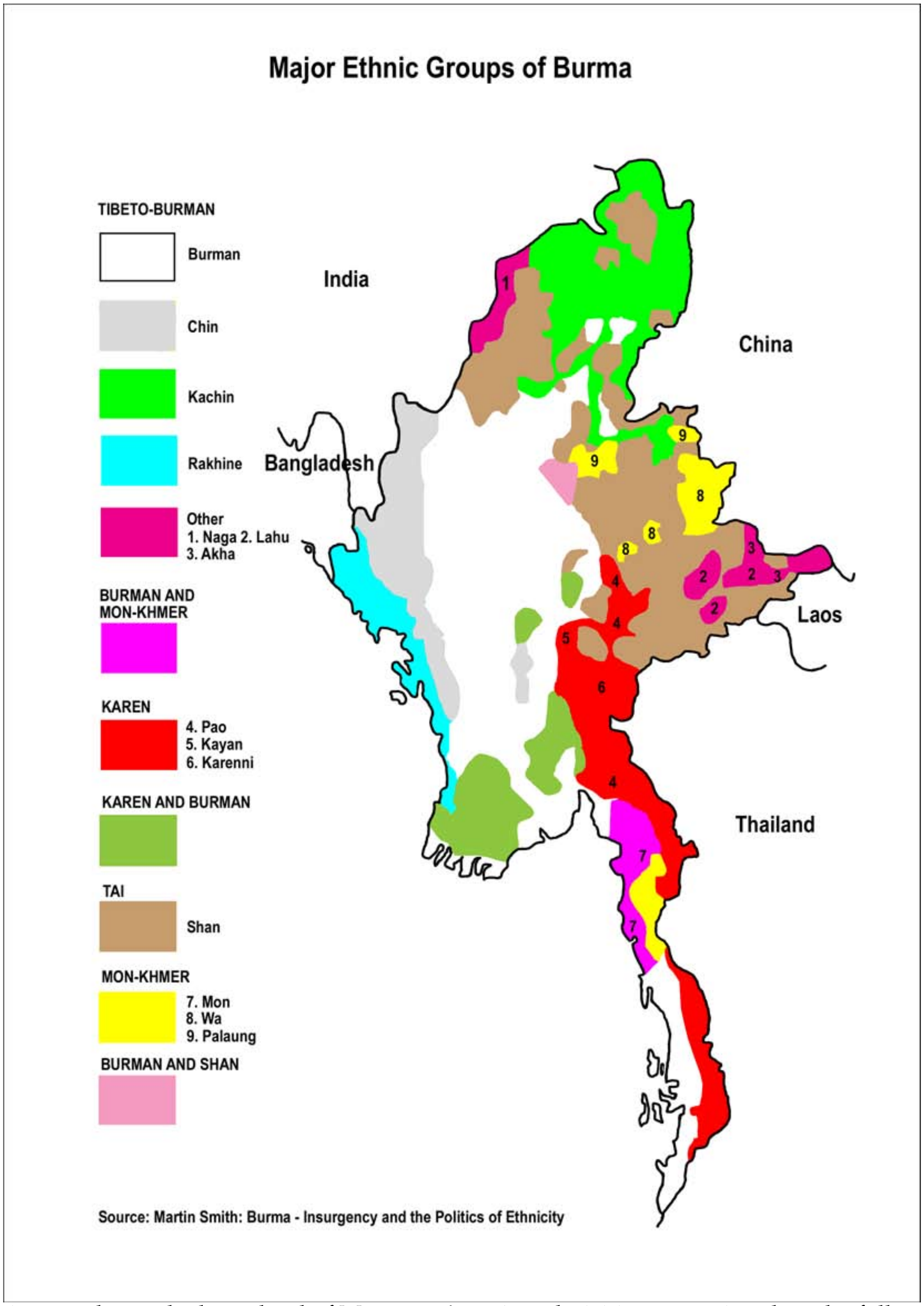
In 2014, there were 459 firms exporting via the land border (see table 3.2). The firm size distribution has a few very large firms which account for a large share of transactions and many small firms. The median firm recorded 12 export transactions in 2014, exported 2 HS6 products to one destination country using one border station. The median shipment value in 2014 was around USD 10,900.

2.2. Ethnic Conflict in Myanmar

Myanmar is a multi-ethnic nation. It counts 135 distinct ethnic groups that are officially recognized (usually assigned to one of eight major ethnic groups[5]). Myanmar's different ethnicities are largely separated geographically. While the Bamar majority (around 68% of the population) mostly lives in the central valleys surrounding the Irrawaddy river, the minorities are clustered in Myanmar's border regions (see the map in figure 3.3).

[5]These eight groups are: Bamar, Chin, Kachin, Kayah, Kayin, Mon, Rakhine and Shan.

Fig. 3.3. Map of Myanmar's Ethnicities



Notes: This map shows the homeland of Myanmar's major ethnicities as reprinted on the following page: `https://reliefweb.int/map/myanmar/myanmar-major-ethnic-groups-burma`. Retrieved on February 17, 2021.

Myanmar's ethnic fractionalization is also reflected in Myanmar's current political system. Ethnic minorities are mostly represented by their own political organisations. These often have an armed wing, the so-called Ethnic Armed Organisations (EAOs). While until the late 2000s ceasefire agreements helped to avoid open confrontations between the military (often called *Tatmadaw*) and EAOs, in the wake of political reforms that opened the country to the outside world ceasefires broke down and open conflict erupted.

Figure 3.4 shows the number of deaths by months recorded by the Uppsala Conflict Data Program (UCDP) from January 2000 to January 2021. As outlined above, the early 2000s were a relatively peaceful period, whereas violence surrounded the heavily contested elections of 2010 and the ensuing democratic transition. The figure also illustrates a large variation from month to month with conflicts suddenly erupting and then cooling down again. This points to the high degree of political uncertainty in Myanmar.

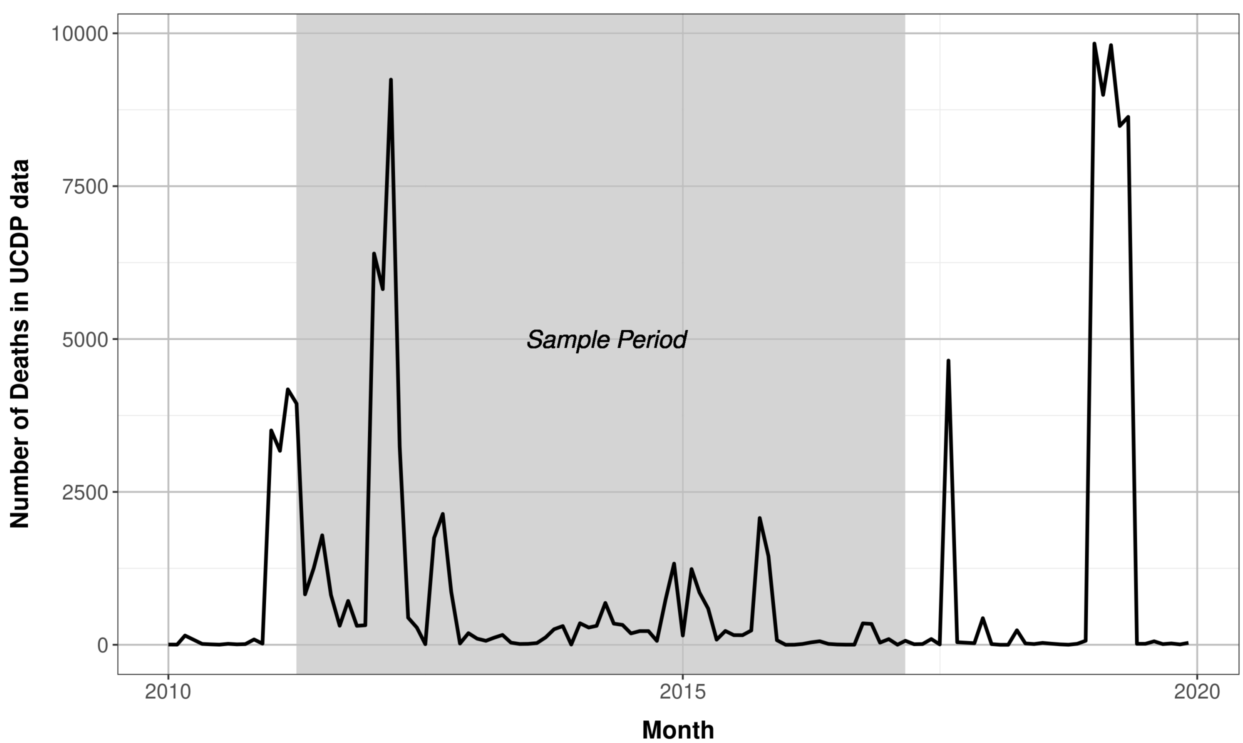


Fig. 3.4. Conflict over Time

Notes: This figure plots the number of deaths recorded by the Uppsala Conflict Data Program between January 2000 and January 2021 in Myanmar. The shaded area corresponds to the sample period used in the analysis below.

While the temporal variation in conflict is high, it is also unevenly distributed across space. Figure 3.1 shows that deadly conflicts are concentrated in Myanmar's borderlands, where

a large part of ethnic minorities live. In the following I will describe three of these localized conflicts that were important during the 2010s and thus my period of analysis.

Northern Kachin state saw a large increase in violence in the early 2010s. In this conflict the *Tatmadaw* and the Kachin Independence Organization (KIO) fought mostly over the control of lucrative jade mines. Christensen et al. (2019) argue that the increased territorial presence in resource-rich areas was driven by the *Tatmadaw*'s fear that the civilian government would tighten control over the distribution of the rents generated by resources.

Also in the early 2010s conflict between the *Tatmadaw* with the Karen National Uninon (KNU) and the DKBA-5 erupted in Karen state which borders Thailand.[6] This conflict, which is part of a decades-long struggle for independence of the Karen people, were ignited by electoral violence surrounding the 2010 elections.

Shan state in the East with the most important trade routes to China has been continuously affected by armed conflict, with a major escalation during the Kokang offensive in February 2015. These were clashes between the *Tatmadaw* and the Myanmar National Democratic Alliance Army (MNDAA), an EAO of ethnic Kokang people, over control of the Kokang self-administered zone (SAZ). The fighting went on until May 2015, resulting in around 200 casualties and over 50,000 displaced civilians. The Kokang SAZ remained under the *Tatmadaw*'s control.[7]

In summary, Myanmar's ethnic conflicts are part of long struggles for independence of its ethnic minorities. These conflicts are often intertwined with disagreements on the distribution of economic rents of which there are plenty in Myanmar's borderlands. Importantly for the analysis below, there is little evidence that conflicts are a function of short-run fluctuations in trade flows. Instead conflicts generates negative spillovers for Myanmar's formal trade.

[6] DKBA-5 (which stands for Democratic Karen Buddhist Army - Brigade 5) is a splinter group of the now dissolved DKBA which was aligned with the *Tatmadaw*.

[7] Figure 3.4 also reveals the unprecedentedly high death toll witnessed during the Rohingya genocide, beginning in October 2016 and entering an even more severe phase in August 2017. These tragic events will not play a major role in this analysis as exports to Bangladesh are only of minor importance when compared with China and Thailand.

3. Data

To study how conflict affects firms' exports, I combine customs transaction data on international trade with geocoded data on conflict.

3.1. *Transactions*

The transaction trade data is collected by Myanmar's Ministry of Commerce. For the period from April 2011 to March 2017 it includes all international trade transactions recorded by customs. Each transaction includes information on the firm, the product (HS10), the value and quantity, the trading partner (country). Crucially, for cross-border trade the data includes the border station at which the transaction was recorded. The transaction level data was then matched to the firm registry (DICA) which allows linking firms to officers and to an address. I then extract the township from this address.

For the period from April 2014 to March 2017, the data includes the exact date of when the transaction was entered into Myanmar's customs registry. For the earlier period the date is entered at the monthly level. Due to this constraint, the analysis below focuses on transactions aggregated to the monthly level.

Time-sensitive Goods. In order to identify goods which are especially affected by disruptions due to conflict, I divide products into seafood and non-seafood based on their HS classification. Seafood is a major export good of Myanmar and the most perishable, as it typically requires a cold chain. It can thus not be stored and delays have a large impact on the value of a shipment. 30% of transactions are in the seafood sector.

To make sure that the effects are not purely driven by factors specific to the seafood sector, I introduce a second way of identifying especially perishable products. In particular, I search the raw product description in the transaction data for the words "alive", "fresh" and "frozen". Then, I compute the share of transactions containing these words at the HS6-level, the definition of a product used below. I classify a product as fresh if the share is above 1%. According to this definition, 25% of transactions are perishable goods. In addition to seafood, this definition contains fresh fruit such as mangoes and also animal products such as meat.

3.2. Routes

I compute the fastest route from each township centroid to each border station. I use the *Open Source Routing Machine* (OSRM) based on *Open Street Map* (OSM) data. My computations are based on the OSM dataset for Myanmar obtained via GeoFabrik on November 1, 2021. Figure 3.1 shows the fastest routes to Myanmar's border stations for a firm located in central Yangon. I then compute the distance between each conflict event and the highways used by the fastest route.

3.3. Conflict

There are two main conflict datasets containing precise geographical information on conflicts used in the literature: *Uppsala Conflict Data Program* (UCDP) and the *Armed Conflict Location and Event Data Project* (ACLED). Besides the the conflict location, both datasets include the actors involved and the number of fatalities associated with a particular conflict. I focus on the former dataset, as the entries in UCDP have been shown to be more accurate and more suitable for subnational analysis of conflict (Eck, 2012). See appendix D for a discussion of the two conflict datasets in the Myanmar context.

For mapping conflicts to roads, it is crucial to measure the event location precisely. The sample is thus restricted to include events where the precision of the information is high.[8] In the period after 2010 this subsample corresponds to 31% of events and 49% of deaths in the UCDP data.

In addition to the choice of dataset, it is unclear which events should be counted when considering the impact on trade flows. Below I mostly consider events that happen within 10km of the highway. In particular, I define conflict as at least five deaths resulting from armed conflict within 10km of a particular route on a given day. This measure picks up large disruptions only, likely leading to underestimating the total effect. Still, in the six-year period of the transaction data (April 2011-March 2017), there are 22 months where access to some border station is blocked for at least some firms.

Figures 3.6 shows the share of affected firm-product-station triplets which are affected by conflict in a given month by border station. We see that a variety of border stations are affected by conflict over the period and conflict is not temporally clustered, i.e. the effects

[8] For UCDP I include events where the locations is measured with precision 1 or 2. The UCDP codebook defines precision level 1 as the exact location of the event and level 2 as the event occurring within a 25 km radius around a known point.

are not identified by one conflict going over several months. Large escalations are rather one-off and short-term events. In addition to these large disruptions, I present results that instead use measures of conflict intensity, using either the number of conflict days in a given months or the total number of conflict-related deaths.

3.4. *Political Connections*

Since some of the actors in Myanmar's conflicts are invested in the export business, this may create strategic interactions between the timing of conflicts and trade flows. It is thus important to account for this firm-level heterogeneity. The mapping of firms to the firm registry which contains names of firm directors allows doing this.

First, the military is an important player both in Myanmar's economy and in its conflicts. Western sanctions initially imposed during the 1990s were targeted at members of the military regime. Governments published lists naming both firms and individuals connected to the military. These can be matched to the transaction data based on the name of the firms, as well as the names of firm directors.[9]

Second, though clearly to a lesser extent than the military, some EAOs may have business interests in the formal economy. Hence, with the help of Burmese natives, I classified firm director names according to Myanmar's main ethnicities.[10] This is largely relying on the fact that many names in Myanmar include ethnicity-specific honorific titles. However, some honorifics are not specific to just one ethnicity and not all Burmese names contain honorifics. I thus relied on multiple native Burmese research assistants for these less clear-cut cases (see appendix C for a more detailed discussion). This procedure allows me to identify firms with non-Bamar directors which I classify as ethnic minority firms.

In the analysis below, I will adopt a broad definition of connectedness to the conflict actors. In particular, all firms with ties to either the military or an ethnic minority are defined as connected. For ethnic minorities this definition will include some firms that have ethnic minority directors but are not necessarily linked to EAOs.[11]

[9]This measure was developed in unpublished work by Felix Forster, Amit Khandelwal, Rocco Macchiavello and Matthieu Teachout which they kindly shared with me. They rely on sanctions lists from the US, the EU and Australia. In addition, they partnered up with a local research company to ensure the accuracy of their matching exercise. Sanctions lists were amended over time and they consider individuals or firms as sanctioned if they ever appear on these lists. Following the previously mentioned authors, state-owned enterprises are also defined as connected to the military.

[10]The main ethnicities are Chin, Shan, Kachin, Kayin, Kayah, Bamar, Rakhine and Mon.

[11]Doing a more precise mapping of firms to EAOs might in theory be possible, but would require

3.5. *Retail Prices*

Myanmar's Central Statistical Office collects monthly data on local retail prices in all of Myanmar's regions. I obtained this data for the period from January 2010 to October 2020. The basket of products changed in September 2016. I was able to identify 116 products present in both sets of products. I manually classify products as tradeable (100 out of 116). In addition, I identify export products when they are present in the raw transaction export data. Table 3.9 contains a list of all products included in the analysis. I end up comparing 29 export products with 71 domestically sold products.

4. Does Conflict Disrupt Trade Flows?

This section shows how conflicts close to trade routes disrupt exports to Myanmar's neighbors and discusses the resulting impact on the local economy.

4.1. *Export Transactions*

The first empirical exercise in this paper tests whether trade flows along a route are lower during a month with armed conflict. To test this I run the following regression

$$\text{Transactions}_{ijpt} = \exp\left[\beta \times \text{Conflict}_{ijt} + \text{FE}_{pd(j)t} + \text{FE}_{ijp}\right] \times \epsilon_{ijpt} \qquad (3.1)$$

where $\text{Transactions}_{ijpt}$ is the number of transactions of firm i via border station j of product p at time t. Conflict_{ijt} captures armed conflict at time t along the route from the location of firm i to border station j. The coefficient of interest is β which we expect to be negative if conflict lowers trade flows.

The specification includes fixed effects at the firm-station-product level, accounting for time-invariant determinants of exports of that firm and product via a particular border station and at the product-destination-time level, controlling for destination-specific demand shocks as well as nationwide supply shocks specific to a certain product.[12] Below I show additional variations where firm-time fixed effects are included, essentially controlling for

substantial support from local research companies which is hard to obtain given the current political circumstances.

[12] Note that one destination country d can have multiple border stations j, but a border station maps to just one destination country.

firm-specific demand and supply shocks.

As is common in the trade literature, I use Poisson Pseudo Maximum Likelihood (PPML) to estimate equation 3.1 to address potential bias stemming from both zeros in the dependent variable and heteroskedasticity (Silva and Tenreyro, 2006). In addition, I present results using ordinary least squares. Standard errors are clustered at the township-station pair, i.e. at the route level.

Table 3.3: Conflict and Transactions

	# Transactions			Any Transaction		
	(1)	(2)	(3)	(4)	(5)	(6)
	Poisson	Poisson	Poisson	OLS	OLS	OLS
Conflict	-0.5082**	-0.4344***	-1.142***	-0.0101***	-0.0154***	-0.0158***
	(0.2095)	(0.1422)	(0.2073)	(0.0031)	(0.0040)	(0.0056)
Observations	727,560	446,194	133,977	727,560	727,560	727,560
R^2	0.25	0.75	0.98	0.13	0.24	0.49
Dependent variable mean	0.44	0.71	2.4	0.06	0.06	0.06
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓			✓		
Month-Product-Destination		✓	✓		✓	✓
Month-Firm			✓			✓

Notes: This table reports the results of estimating equation 3.1. Standard errors are clustered at the route level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Table 3.3 presents the results of estimating equation 3.1. Conflict has a large impact on the number of transactions passing through the affected route. The coefficient in the second column shows that a major conflict leads to an average drop in transactions of 43%. The extensive margin results using OLS presented in the last three columns of the table confirm a large and robust effect of conflict on export flows. Column 5 indicates that during a month with conflict the probability of observing a conflict drops by 1.5 percentage points.

The inclusion of a firm-time fixed effect roughly doubles the PPML estimate. The negative effect of conflict on exports is thus not driven by conflict affecting the production of firms directly. A larger coefficient (in absolute terms) may be driven by substitution across export routes by firms, as now identification comes from changes in the relative use of export

routes by the same firm, depending on whether these routes are currently affected by conflict, whereas previously identification was coming from the differential use of export routes to the same destination for one product, both across and within firms. Note that the inclusion of firm-time effects excludes firms that export only using one border city and thus the effective sample reduces to a few, large firms.

Endogeneity Concerns. In order to obtain an unbiased estimator when estimating equation 3.1, the expectation of the error term conditional on regressors needs to be zero. The most relevant threat to this assumption in this paper is reverse causality. In particular, more trade along a given route could raise the gains of controlling it, e.g. because of taxation, and hence the likelihood of observing conflict.

First, notice that this reverse causality would likely lead to an upward bias in the estimated coefficient. Thus, if anything, the results presented here would underestimate the true negative effect of conflict on trade. Nonetheless, I argue below that this mechanism is unlikely to be at play, at least in the short time window that I am considering here.[13]

Second, the main definition of conflicts I use happens to only include killings of civilians by the military because these typically result in a larger number of deaths than standoffs between EAOs and the military. The nature of these conflicts is thus based on longer standing conflicts between ethnic minorities and the military. As described in section 2, the military often wants to drive ethnic minorities from resource-rich pieces of land. The motives behind these atrocities are thus not driven by short-run fluctuations in trade.

Third, if taxing goods passing through were a driver of conflict, we would expect a different coefficient for more valuable goods, as they would likely result in higher tax revenue. I compute the average price and shipment value for each product and then estimate equation 3.1 interacting the conflict variable with these values. The results presented in table 3.10 show that there is no differential effect for higher value products.

Some of the parties to the conflict are closely linked to trading firms. In particular, the military is both present in many sectors of the economy as well as actor in the majority of conflicts. Another threat to identification could thus come from the fact that the parties involved in the conflict simultaneously decide on when to do business and when to engage

[13] In general, control of trade routes in the long run is a likely driver of conflict. Rebel groups sometimes generate a substantial share of their income from taxing trade. Studying how increases in trade can cause turf wars is an interesting path for future research. Dell (2015) shows that drug trade can increase violence along trafficking routes in Mexico.

in heavy fighting. Reassuringly, when I drop firms directly connected to the conflict parties, I find very similar coefficients (see column 4 of table 3.11). As I show below, if anything, firms connected to the military actually export more during conflict periods.

In addition, I use a lagged version of the explanatory variable. If trade and conflict are determined simultaneously, lagging the explanatory alleviates this concern. Column 3 of table 3.11 shows that the lagged conflict variable leads to a very similar estimate, indicating that conflict is not strategically halted when trade occurs, but rather that it creates an environment of insecurity where firms restrict their exports even one month later.

To add to this, I estimate a fully dynamic specification.[14] The results in figure 3.7 show that there is no anticipation of conflicts and that they still have an effect in the following period. Finally, there is some evidence of a catching-up three months later. In the case of attacks coordinated with firms, one would expect some frontloading of transactions, which does not seem to be the case in the data.

Conflict Definition. To further ensure the validity of these results, I use different definitions of conflict (see table 3.12). First, I use a different conflict dataset (ACLED).[15] I continue to find a strong and precisely estimated negative effect of conflict on trade. The magnitude of the coefficient in column 2 is lower which is in line with the ACLED data including a wider set of conflict events which comes at the cost of introducing measurement error.

Second, I use two measures of conflict intensity, the number of conflict days and the total number of deaths related to these conflicts. Columns 4 and 5 show that both of these measures result in a precisely estimated negative coefficient. Column 5 suggests, for example, that a 10% increase in conflict deaths decreases exports by 0.6%.

Aggregation. So far the analysis has been at the firm-route-product level. To further underline the aggregate implications of conflict, I instead estimate equation 3.1 at the destination country-product level. I focus on disruptions along the most important trade route between Yangon and Muse. The highway section between Mandalay and Muse is both a crucial artery for trade as well as frequently disrupted by conflict in Shan State.

[14] To do this, I create a full vector of observations for each firm-product-route and treatment date. I then estimate coefficients relative to two rather than one baseline period, as recommended by Borusyak and Jaravel (2017) to avoid multicollinearity. Note that other recent methods as developed by De Chaisemartin and d'Haultfoeuille (2020) are not suitable when using a non-linear PPML estimator.

[15] See section 3 for a discussion on conflict data

Table 3.13 shows that conflict along this highway has an overall negative impact on export transactions to China. In the first column of this table I focus on exports to China and rely purely on the variation across time, controlling for seasonality of exports. The second column includes exports to other countries, which allows controlling for time fixed effects and thus exploiting variation across destination countries. This implies that localized disruptions lead to more aggregate effects and substitution across routes or firms does not offset the negative effect caused by the disruptions.

4.2. Pass-through to Local Prices

To further study the aggregate impact of the disruptions to market access, I look at how local prices for export goods change during conflict periods. Anecdotal evidence suggests that firms who usually export are forced to sell part of their production domestically when there is conflict along the routes to neighboring countries. To test this, I use the data on retail prices. I focus on the Yangon region which is itself not affected by armed conflict and is the main trade hub of Myanmar. I run the following regression

$$\log(\text{price}_{pt}) = \beta \times \text{Conflict}_t \times \text{Export}_p + \text{FE}_p + \text{FE}_t + \epsilon_{pt} \tag{3.2}$$

where price_{pt} is the retail price of product p in month t in the Yangon region, Conflict_t captures armed conflict along the major trade route to China (the Mandalay-Muse highway) and Export_p is a dummy indicating whether product p is exported to China.[16] The specification includes product and time fixed effects. Below I also estimate equation 3.2 with product-calendar month fixed effects, accounting for seasonality patterns and product-specific linear trends. The sample consists of all tradeable products in the retail price data, but I present additional results where I include non-tradeables. Standard errors are clustered at the product level.

If some of the production destined for export has to be sold on the domestic market, the coefficient of interest β should be negative. As table 3.4 shows, this is indeed the case. When there is conflict along the major highway to China, the retail price in Yangon of export goods relative to goods only sold domestically drops by around 5%. This suggests that at least some producers choose to sell on the domestic market rather than export when

[16] I focus on the Mandalay-Muse highway to keep the analysis simple. As table 3.1 shows, Muse is, by far, the most important border city and alone accounts for more than one billion US dollars in exports.

there is conflict along the Mandalay-Muse highway, resulting in potentially substantial economic losses.

Table 3.4: Pass-Through to Local Prices

	Retail Price in Yangon (logs)			
Period:	1/2010-10/2020			04/2011-03/2017
	(1)	(2)	(3)	(4)
Conflict M-M Highway × Export Good	-0.0532**	-0.0494*	-0.0418***	-0.0549**
	(0.0238)	(0.0251)	(0.0152)	(0.0224)
Observations	11,035	11,035	11,035	6,864
R^2	0.97	0.97	0.99	0.99
Fixed Effects				
Product	✓			
Month	✓	✓	✓	✓
Calendar Month-Product		✓	✓	✓
Linear Trend				
Month × Product			✓	✓

Notes: This table reports the results of estimating equation 3.2. Standard errors are clustered at the product level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Robustness. The result is robust to the inclusion of calendar month-product fixed effects, accounting for seasonality. In addition, including a product-level linear trend does not change the result, capturing any long-run price trends at the product level that could be correlated with conflict incidence. Finally, restricting the estimation window to the period of the trade transaction data does not change the main finding that prices of export products drop in times of closed trade routes.

Finally, I have so far omitted non-tradeables from the control group of non-exported goods. As shown in table 3.14, including them increases the estimated magnitude. This might be due to a higher degree of substitutability among tradeable goods, compared to tradeable versus non-tradeables.

4.3. Discussion

These results suggest that conflicts have a substantial direct negative effect on trade flows. Taking the probability of a firm-product-station triplet being affected by conflict in the full sample (3.2%), the results imply that in this setting conflict reduces export transactions by around 1.4% over the sample period. If we use the alternative conflict data from ACLED, the much higher baseline probability (19.5%) leads to an estimated reduction in export transactions of 5.4%. This captures only the largest conflicts and their direct impact. The total effect of conflict is likely larger.

The results on retail prices imply that traders lose access to foreign markets and switch to selling domestically, indicating that withstanding disruptions and continuing to export is difficult. Prices drop by 5% when the major trade route with China is disrupted, again suggesting substantial losses, potentially spilling over to purely domestically active firms. This finding demonstrates the interdependence between exports and the domestic markets. In canonical trade models this interdependence is typically assumed away and firms decide whether to enter and how much to sell independently for each market. In a recent exception Almunia et al. (2021) describe how firms compensate a negative domestic demand shock by exporting more, a phenomenon the authors label *venting out*. I add to this nascent literature by showing evidence for firms *venting in* in the face of disruptions to foreign market access.

Overall, the results suggest that conflicts have an economic impact in locations that are far away from where violent conflict occurs. This adds to the literature that tries to estimate the economic impact of war and conflict. While Korovkin and Makarin (2021) and Couttenier et al. (2022) show that conflict can disrupt supply chains and thereby lead to spillovers along the production network, I add to this by providing evidence that conflict and insecurity can substantially increase trade costs and hurt trade relationships.

The question arises why exporters cannot switch export routes and simply continue trading. The specific nature of trade relationships in this context makes this difficult. First, the market structure is such that coordinating ex ante with buyers is not common. Many of the agricultural commodities are actually sold on the spot. An exporter transports the goods to one of the border cities. Then, she sells her goods in a marketplace to buyers that cross the border from the neighboring country. This market structure makes it difficult to contract ex-ante and to deliver the goods via a different border city.

Second, the road network in Myanmar is sparse. Reaching a border city via a different route and thereby circumventing the conflict-affected section is often not possible. The few roads that cross Myanmar's mountainous borderlands, such as the Mandalay-Muse highway, are often crucial infrastructure bottlenecks.

Third, for many of these products using different transport modes is difficult. Switching e.g. to maritime trade requires complying with different regulations and would typically require interacting with different buyers since at least the buyers on the Chinese border are located far away from the next port. Only a small fraction of firms is active both in maritime and land border trade.

5. Using Perishability to Identify Risk

In order to identify how the risk of exporting during conflict periods affects firms and products differentially, I rely on product characteristics. In particular, I separate the seafood sector from other products. Seafood is highly perishable and due to the lack of a reliable cold chain subject to high losses when shipments are delayed. First, I show that exporters can generally substitute shipments intertemporally, but not in the seafood sector. Second, I look at what types of firms are more resilient to the increased risk, using the seafood sector to identify highly risky transactions.

5.1. *Intertemporal Substitution*

Disruptions of trade routes caused by conflict are only temporary and typically last a few days. This means exporters can potentially delay earlier shipments and lump them together with later ones. If there is intertemporal substitution in demand and producers can store products, we would thus expect exporters to adjust upwards the shipment size once trade routes are open again to compensate for the missing transactions. To test this I run the following specification

$$\log(\text{Transaction Size}_{ijpt}) = \beta \times \text{Conflict}_{ijt} + \text{FE}_{pd(j)t} + \text{FE}_{ijp} + \epsilon_{ijpt} \tag{3.3}$$

where $\text{Transaction Size}_{ijpt}$ is the average value of a transaction by firm i via border station j of product p at time t. The explanatory variables and the clustering of standard errors are as in equation 3.1. The estimation is done using ordinary least squares. Since the average

transaction size is highly skewed, I use the log of transaction size as dependent variable.

Table 3.5 shows that transaction size increases by around 20% during months with conflict. While exporters decrease the amount of export transactions the average size increases, compensating for the disruption. As expected, the intertemporal substituability of trade transactions strongly depends on product characteristics. In particular, perishable goods can only be stored for a certain amount of time, potentially subject to high costs, e.g. refrigeration. Accordingly, I find that only producers of non-seafood exports increase average transaction size and are thus able to shift their transactions across time. In contrast, exporters of seafood see no increase in transaction size. This suggests that delaying transactions is a feasible adjustment strategy for less time-sensitive products, limiting the economic cost from disruptions.

Table 3.15 shows that the results are robust to using a linear dependent variable and table 3.16 uses a different definition of perishability (see section 3).

Table 3.5: Conflict and Transaction Size

	Avg. Transaction Value (logs)					
	Full Sample	Seafood	Non-Seafood	Full Sample	Seafood	Non-Seafood
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	0.1011**	0.0473	0.1461***	0.2047***	0.0567	0.2279***
	(0.0402)	(0.0567)	(0.0552)	(0.0670)	(0.1530)	(0.0790)
Observations	42,033	10,592	31,441	42,033	10,592	31,441
R^2	0.85	0.83	0.85	0.92	0.91	0.92
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓	✓	✓			
Month-Product-Destination				✓	✓	✓

Notes: This table reports the results of estimating equation 3.3. Standard errors are clustered at the route level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

5.2. Firm Heterogeneity

In order to study heterogeneity across firms, I estimate the following variation of equation 3.1

$$\text{Transactions}_{ijpt} = \exp\left[\beta \times \text{Conflict}_{ijt} + \gamma \times \text{Conflict}_{ijt} \times X_i + \text{FE}_{pd(j)t} + \text{FE}_{ijp}\right] \times \epsilon_{ijpt} \tag{3.4}$$

where X_i is a characteristic of the firm. Other variable definitions and the estimation strategy is as in equation 3.1. To study decision-making under risk, I divide the sample into seafood and other exports. As the previous analysis has shown, firms cannot intertemporally substitute shipments of seafood products, making handling them riskier for firms. We thus expect a larger differential between more and less risk-averse firms in the seafood sector relative to other products.

5.2.1. Firm Size

First, I look at whether firm size is related to how firms deal with risk. Large firms have many potential advantages during conflict periods. For example, they may face a lower probability of bankruptcy when losing a shipment. They can thus continue exporting even when it becomes risky. Small and financially constraint firms no longer find it optimal to trade.

I thus estimate equation 3.4, defining X_i as firm size. I measure firm size by the value of all transactions in non-conflict periods and then define a large firm as being in the top quartile. Table 3.6 shows the results. In general, large firms are more likely to continue exporting during conflict periods. The differential between small and large firms is especially pronounced in the seafood sector, with the large firm premium being at least twice as large.

Table 3.6: Conflict and Firm Size

	# Transactions					
	Seafood			Non-Seafood		
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	-0.4686**	-1.581***	-1.514***	-0.4156*	-0.5192**	-0.7741***
	(0.2284)	(0.4300)	(0.4814)	(0.2174)	(0.2171)	(0.1667)
Conflict × Large Firm		1.159***	1.068**		0.1282	0.4111**
		(0.4411)	(0.4498)		(0.1918)	(0.1667)
Observations	156,456	156,456	109,531	571,104	571,104	336,663
R^2	0.19	0.19	0.31	0.32	0.32	0.78
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓	✓		✓	✓	
Month-Destination-Product			✓			✓

Notes: This table reports the results of estimating equation 3.4 with X_i defined as firm size. Standard errors are clustered at the route level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

Robustness. The results are robust to different ways of defining firm size. First, I use the value of firm exports until the end of 2012 to define initial firm size and then use only transactions in the estimation sample from after 2012. This approach has the disadvantage of reducing the sample size substantially since there was substantial export entry during the period of analysis. Table 3.17 shows the results are nevertheless robust to this reduced estimation sample.

Second, I measure firm size continuously. Table 3.17 shows that the sign of the interaction between conflict and (log) firm size is positive and statistically significant and thus in line with the previous results.

Third, I use a different definition of the perishable sector (see section 3 for a description). Table 3.19 shows that the firm size differential is again larger for the more perishable products.

5.2.2. *Connections to the Military*

While the military was formally not in power during the sample period, political connections to the military are still believed to have been important. The *Tatmadaw* are a conflict party in a vast majority of conflicts. In order to test whether connected firms enjoy better protection from conflict risk, I estimate equation 3.4 defining X_i as connected. The results in table 3.7 show that connected seafood exporters are more resilient during conflict periods than their non-connected counterparts. In contrast, there is no statistically significant difference in the effect for non-seafood exporters. This result is robust to controlling for a time-varying effect of being connected to the military. This is important in case the probability of conflict is correlated with general political turmoil that affects connected firms differentially. In addition, I show that using a different definition of perishable products does not qualitatively change the results (see table 3.20).

Table 3.7: Conflict and Political Connections

	# Transactions					
	Seafood			Non-Seafood		
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	-0.5253**	-0.5265	-0.5205**	-0.4109*	-0.4102***	-0.4117*
	(0.2438)	(0.3814)	(0.2444)	(0.2187)	(0.1283)	(0.2189)
Conflict × Connected	0.7423***	0.6601*	0.5773*	-0.2781	-0.3678	-0.5788
	(0.2357)	(0.3380)	(0.3459)	(0.6739)	(0.6228)	(1.101)
Observations	156,456	109,531	156,224	571,104	336,663	570,509
R^2	0.19	0.31	0.19	0.32	0.78	0.32
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓			✓		
Month-Destination-Product		✓			✓	
Month-Connected			✓			✓

Notes: This table reports the results of estimating equation 3.4 with X_i defined as a dummy variable indicating whether a firm is connected to the military. Standard errors are clustered at the route level. ***, **, * denote statistical significance at the 1%, 5% and 10% levels, respectively.

5.3. *Discussion*

The analysis shows that exporting seafood is risky because of its perishability. In contrast to less perishable products, it does not allow for intertemporal substitution. Expectedly, this suggests that stockpiling is a major safeguard against short-run disruptions of market access. While there is evidence that stockpiling inputs can shield firms from global supply chain disruptions (Lafrogne-Joussier et al., 2022), this analysis shows that suppliers can also make use of this and shift transactions into the future.

I show that large firms are more resilient during periods of increased risk, especially in the more risky seafood sector. This suggests that trade cost uncertainty may skew the firm size distribution towards larger firms. In line with this I show in figure 3.8 that seafood accounts for a larger share of exports in the upper tail of the firm size distribution.

In addition, I show that firms connected to the military may enjoy better protection and can continue to export during conflict periods. This finding highlights how being connected to powerful actors can become more important during periods of turmoil. To the best of my knowledge this finding is new to the literature on how firms can benefit from political connections in developing countries.[17]

6. Conclusion

This paper studies the effects of sudden, unanticipated disruptions of trade routes on exports in a developing country context. In particular, the analysis reveals that ethnic conflicts along trade routes in Myanmar lead firms to decrease their exports along these routes. Especially, exporters of perishable and thus time-sensitive products are affected which face potentially high losses when their shipments are delayed. Local prices for export goods drop, suggesting that producers sell some of their produce on domestic markets. Firms with connections to the military and large exporters are the most resilient and are more likely than their competitors to keep exporting perishable, more risky products.

Understanding the challenges faced by exporters in developing countries is key to designing policies that increase participation in international trade networks. In Myanmar, localized disruptions to road networks can have large consequences because firms are not sufficiently diversified in terms of markets they serve. Reducing entry barriers to foreign

[17] In a recent survey of the literature Atkin et al. (2022) point to connections to the military as an area unexplored in the literature on the nexus of trade and development.

markets and a strong domestic market allows firms to shift sales and become more resilient. In addition, a denser road network could reduce the importance of particular bottlenecks.

For the economic outlook of Myanmar these results are troubling. If the military regime stays in power, Western sanctions will further increase the country's dependence on trade with China. At the same time, after the coup d'état, conflict between the *Tatmadaw* and EAOs has increased in the Burmese border regions. In addition to the human tragedy this causes, levels of trade will remain low for some time to come and impede the economic development of Myanmar.[18]

[18] In fact, recent numbers suggest that exports via the land border have dropped significantly. This is likely due in part to the ongoing conflicts in Myanmar's border regions (see e.g. `https://www.thestar.com.my/aseanplus/aseanplus-news/2022/02/24/myanmar039s-sea-trade-surges-in-interim-budget-period`).

References

Almunia, M., Antràs, P., Lopez-Rodriguez, D., and Morales, E. (2021). Venting out: Exports during a domestic slump. *American Economic Review*, 111(11):3611–62.

Atkin, D. and Donaldson, D. (2015). Who's getting globalized? the size and implications of intra-national trade costs. Technical report, National Bureau of Economic Research.

Atkin, D., Khandelwal, A., Boudreau, L., Dix-Carneiro, R., Isabela, M., Medina, P., McCaig, B., Morjaria, A., Pascali, L., Rijkers, B., and Startz, M. (2022). International trade. *VoxDevLit*, 4(1).

Barrot, J.-N. and Sauvagnat, J. (2016). Input specificity and the propagation of idiosyncratic shocks in production networks. *The Quarterly Journal of Economics*, 131(3):1543–1592.

Boehm, C. E., Flaaen, A., and Pandalai-Nayar, N. (2019). Input linkages and the transmission of shocks: Firm-level evidence from the 2011 tōhoku earthquake. *Review of Economics and Statistics*, 101(1):60–75.

Borusyak, K. and Jaravel, X. (2017). Revisiting event study designs. *Available at SSRN 2826228*.

Chiovelli, G., Michalopoulos, S., and Papaioannou, E. (2018). Landmines and spatial development. Technical report, National Bureau of Economic Research.

Christensen, D., Nguyen, M., and Sexton, R. (2019). Strategic violence during democratization: Evidence from myanmar. *World Politics*, 71(2):332–366.

Couttenier, M., Monnet, N., and Piemontese, L. (2022). The economic costs of conflict: A production network approach. Discussion Paper 16984, Centre for Economic Policy Research.

De Chaisemartin, C. and d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Dell, M. (2015). Trafficking networks and the mexican drug war. *American Economic Review*, 105(6):1738–79.

Eck, K. (2012). In data we trust? a comparison of ucdp ged and acled conflict events datasets. *Cooperation and Conflict*, 47(1):124–141.

Ferguson, J. M. (2015). Who's counting?: Ethnicity, belonging, and the national census in burma/myanmar. *Bijdragen tot de taal-, land-en volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia*, 171(1):1–28.

Handley, K. and Limão, N. (2017). Policy uncertainty, trade, and welfare: Theory and evidence for china and the united states. *American Economic Review*, 107(9):2731–83.

Handley, K. and Limão, N. (2022). Trade policy uncertainty. Working Paper 29672, National Bureau of Economic Research.

Korovkin, V. and Makarin, A. (2021). Production networks and war. *Available at SSRN*.

Ksoll, C., Macchiavello, R., and Morjaria, A. (2021). Electoral violence and supply chain disruptions in kenya's floriculture industry. Technical report, National Bureau of Economic Research.

Lafrogne-Joussier, R., Martin, J., and Mejean, I. (2021). *Supply shocks in supply chains: Evidence from the early lockdown in China*. Centre for Economic Policy Research.

Lafrogne-Joussier, R., Martin, J., and Mejean, I. (2022). Supply shocks in supply chains: Evidence from the early lockdown in china. *IMF economic review*, pages 1–46.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and statistics*, 88(4):641–658.

Appendix

A. Additional Tables

Table 3.8: Myanmar's Main Exports via Land Border in 2014

HS Code	Product Description	Transactions	Value (mln USD)
0302	Fish; fresh or chilled, excluding fish fillets and other fish meat of heading 0304	6109	83.40
0301	Fish; live	5917	50.12
1006	Rice	4422	401.95
0306	Crustaceans; in shell or not, live, fresh, chilled, frozen, dried, salted or in brine; smoked, cooked or not before or during smoking; in shell, steamed or boiled, whether or not chilled, frozen, dried, salted or in brine; edible flours, meals, pellets	3725	76.70
0807	Melons (including watermelons) and papaws (papayas); fresh	2134	100.58
1005	Maize (corn)	2050	285.41
0713	Vegetables, leguminous; shelled, whether or not skinned or split, dried	1937	102.25
0802	Nuts (excluding coconuts, Brazils and cashew nuts); fresh or dried, whether or not shelled or peeled	1327	26.31
0305	Fish, dried, salted or in brine; smoked fish, whether or not cooked before or during the smoking process; flours, meals and pellets of fish, fit for human consumption	1103	14.63
1211	Plants and parts of plants (including seeds and fruits), used primarily in perfumery, pharmacy; for insecticidal, fungicidal or similar purposes, fresh or dried, whether or not crushed or powdered	950	8.84

Table 3.9: List of Products in Retail Price Data

Export Good	Non-Export Good	Non-Export Good (cont.)	Non-Export Good (cont.)
Banana	Bicycle	Gourd	Taste powder (50gm)
Beef	Bread	Groundnut Oil	Tetron
Beer (Mandalay)	Bricks	Horlicks (1lb)	Thanakha
Betel leaves	Brinjal	Long bean	Timber(Hard wood)
Betel nut	Bulb	Long cloth(white)	Tooth paste(Pepsodent)
Cabbages	Candles	Milk powder (2 1/2lb)	Towel
Cane jaggery	Casette-Radio	Mutton	Umbrella(Foreign)
Coconut	Cement(50 kg)	Nail	Vegetable Oil
Condensed Milk (local)	Ceramic Ware Bowl	Ngabokechauk	Vest(1/30 cord)
Dried chillies (short)	Ceramic Ware Plate	Ngagyi	Washing soap(Shwewa)
Fish ngapi	Charcoal	Ngagyichauk	Wheat Flour
Garlic (single)	Cheroots	Ngagyin	Woman's cloth
Gram	Chicken	Ngakhu	Zinc(32 gate)
Lemon	Chicken Eggs	Ngakunshutchauk	
Lime	Cigarette (Duya)	Ngayan	
Ngamoke	Cigarette (Foreign)	Ngayanchauk	
Ngamyitchin	Coconut oil	Ovaltine (1lb)	
Onions (big)	Coffee mix / Tea mix	Palm Oil	
Papaya	Condensed Milk (Foreign)	Palm jaggery	
Pegyi	Dhani(Laputta)	Penilay	
Pork	Diecel	Petrol	
Potatoes	Dried Prawns	Radish	
Rice(Emata)	Duck	Salt	
Roselle leaf	Duck Eggs	Salted fish	
Sadawpe	Electric Iron	Shampoo (Tayaw)	
Seamum Oil	Exercise book,80 pages	Shrimp nganpyaye	
Sugar	Fire wood	Slipper(Foreign)	
Tomatoes	Fish nganpyaye	Slipper(Leather)	
Vermicelli (rice)	Fresh Milk	Sport-shirt	

Table 3.10: Conflict and Transactions - Heterogeneity by Value

	# Transactions				
	(1)	(2)	(3)	(4)	(5)
Conflict	-0.4344***	-0.4331	-0.4334***	-0.4148**	-0.4974***
	(0.1422)	(0.5623)	(0.1612)	(0.1662)	(0.1547)
Conflict × High Transaction Value		-0.0014			
		(0.5576)			
Conflict × Transaction Value (HS6 Average)			-1.33×10^{-8}		
			(4.27×10^{-7})		
Conflict × High Price				-0.0326	
				(0.2391)	
Conflict × Price (HS6 Average)					5.02×10^{-5}
					(0.0001)
Observations	446,194	446,194	446,194	446,194	446,194
R^2	0.75	0.75	0.75	0.75	0.75
Fixed Effects					
Month-Destination-Product	✓	✓	✓	✓	✓
Firm-Route-Product	✓	✓	✓	✓	✓

Table 3.11: Conflict and Transactions - Robustness

	# Transactions		
	Baseline	Drop Connected	Lagged
	(1)	(2)	(3)
Conflict	-0.4344***	-0.5441***	
	(0.1422)	(0.1367)	
Conflict (lagged)			-0.3966**
			(0.1549)
Observations	446,194	343,032	435,879
R^2	0.75	0.61	0.75
Fixed Effects			
Firm-Route-Product	✓	✓	✓
Month-Product-Destination	✓	✓	✓

Table 3.12: Conflict and Transactions - Different Conflict Definitions

	# Transactions				
	Baseline	ACLED	ACLED (Civilians)	Conflict Intensity	
	(1)	(2)	(3)	(4)	(5)
Conflict	-0.4344***				
	(0.1422)				
Conflict (ACLED)		-0.2778**			
		(0.1148)			
Conflict (ACLED, Civilians)			-0.4279**		
			(0.2066)		
IHS(Conflict Days)				-0.1914***	
				(0.0649)	
IHS(Deaths in 50km)					-0.0634**
					(0.0294)
Observations	446,194	446,194	446,194	446,194	446,194
R^2	0.75	0.74	0.75	0.75	0.75
Fixed Effects					
Firm-Route-Product	✓	✓	✓	✓	✓
Month-Product-Destination	✓	✓	✓	✓	✓

Table 3.13: Conflict and Transactions - Aggregate Impact

	# Transactions	
	China Only	All Destinations
	(1)	(2)
Conflict M-M Highway	-0.4927***	-0.3868**
	(0.1644)	(0.1668)
Observations	24,552	67,392
R^2	0.20	0.43
Fixed Effects		
Destination-Product	✓	✓
Calendar Month	✓	
Month		✓

Table 3.14: Retail Prices and Conflict - Including Non-tradeables

	Retail Price in Yangon (logs)			
Period:	1/2010-10/2020			04/2011-03/2017
	(1)	(2)	(3)	(4)
Conflict along M-M Highway × Export Good	-0.0915***	-0.0934***	-0.0609***	-0.0810***
	(0.0296)	(0.0320)	(0.0161)	(0.0238)
Observations	12,915	12,915	12,915	7,988
R^2	0.94	0.94	0.98	0.99
Fixed Effects				
Product	✓			
Month	✓	✓	✓	✓
Calendar Month-Product		✓	✓	✓
Month × Product			✓	✓

Table 3.15: Conflict and Transaction Size - Linear Dependent Variable

	Avg. Transaction Value					
	Full Sample	Seafood	Non-Seafood	Full Sample	Seafood	Non-Seafood
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	12,436.2*	1,381.9	17,631.5**	24,632.1***	-3,176.7	28,990.1***
	(6,587.8)	(1,797.1)	(8,828.7)	(7,025.6)	(6,994.4)	(8,122.5)
Observations	42,033	10,592	31,441	42,033	10,592	31,441
R^2	0.69	0.71	0.69	0.88	0.82	0.88
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓	✓	✓			
Month-Product-Destination				✓	✓	✓

Table 3.16: Conflict and Transaction Size - Different Split

	Avg. Transaction Value (logs)					
	Full Sample	Fresh	Non-Fresh	Full Sample	Fresh	Non-Fresh
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	0.1011**	0.0599	0.1194**	0.2047***	0.0436	0.2311***
	(0.0402)	(0.0623)	(0.0485)	(0.0670)	(0.1498)	(0.0751)
Observations	42,033	9,204	32,829	42,033	9,204	32,829
R^2	0.85	0.84	0.85	0.92	0.93	0.92
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓	✓	✓			
Month-Product-Destination				✓	✓	✓

Table 3.17: Conflict and Initial Firm Size

	# Transactions			
	Seafood		Non-Seafood	
	(1)	(2)	(3)	(4)
Conflict	-1.088**	-1.417**	-0.1404	-0.4428**
	(0.4356)	(0.5563)	(0.1621)	(0.1834)
Conflict × Large Firm	1.199**	1.134***	0.5200***	0.6804***
	(0.4720)	(0.3454)	(0.1878)	(0.2019)
Observations	46,818	32,808	103,377	57,507
R^2	0.28	0.42	0.40	0.80
Fixed Effects				
Firm-Route-Product	✓	✓	✓	✓
Month	✓		✓	
Month-Destination-Product		✓		✓

Table 3.18: Conflict and Continuous Firm Size

	# Transactions			
	Seafood		Non-Seafood	
	(1)	(2)	(3)	(4)
Conflict	-4.084**	-4.863***	-3.582***	-4.476***
	(1.776)	(1.856)	(1.146)	(0.9236)
Conflict × log(Firm Size)	0.2132**	0.2588**	0.1897***	0.2390***
	(0.1003)	(0.1063)	(0.0707)	(0.0553)
Observations	156,384	109,459	570,816	336,439
R^2	0.19	0.31	0.32	0.78
Fixed Effects				
Firm-Route-Product	✓	✓	✓	✓
Month	✓		✓	
Month-Destination-Product		✓		✓

Table 3.19: Conflict and Firm Size - Different Split

	# Transactions					
	Fresh			Non-Fresh		
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	-0.6847***	-1.769***	-1.243***	-0.4574**	-0.6247***	-0.8313***
	(0.2640)	(0.4221)	(0.4302)	(0.2275)	(0.2290)	(0.1684)
Conflict × Large Firm		1.145***	0.7504**		0.2013	0.4897***
		(0.4270)	(0.3710)		(0.1933)	(0.1692)
Observations	147,672	147,672	88,650	579,888	579,888	357,544
R^2	0.32	0.32	0.90	0.37	0.37	0.69
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓	✓		✓	✓	
Month-Destination-Product			✓			✓

Table 3.20: Conflict and Political Connections - Different Split

	# Transactions					
		Fresh			Non-Fresh	
	(1)	(2)	(3)	(4)	(5)	(6)
Conflict	-0.7644***	-0.5733	-0.7619***	-0.4520**	-0.3970***	-0.4516**
	(0.2795)	(0.3622)	(0.2849)	(0.2299)	(0.1302)	(0.2304)
Conflict × Connected	1.013***	0.7008**	0.8888**	-0.2683	-0.2467	-0.5496
	(0.3158)	(0.3568)	(0.4032)	(0.6306)	(0.5758)	(0.7735)
Observations	147,672	88,650	147,588	579,888	357,544	579,753
R^2	0.32	0.90	0.34	0.37	0.69	0.37
Fixed Effects						
Firm-Route-Product	✓	✓	✓	✓	✓	✓
Month	✓			✓		
Month-Destination-Product		✓			✓	
Month-Connected			✓			✓

B. Additional Figures

Fig. 3.5. Map of Myanmar's Exporters

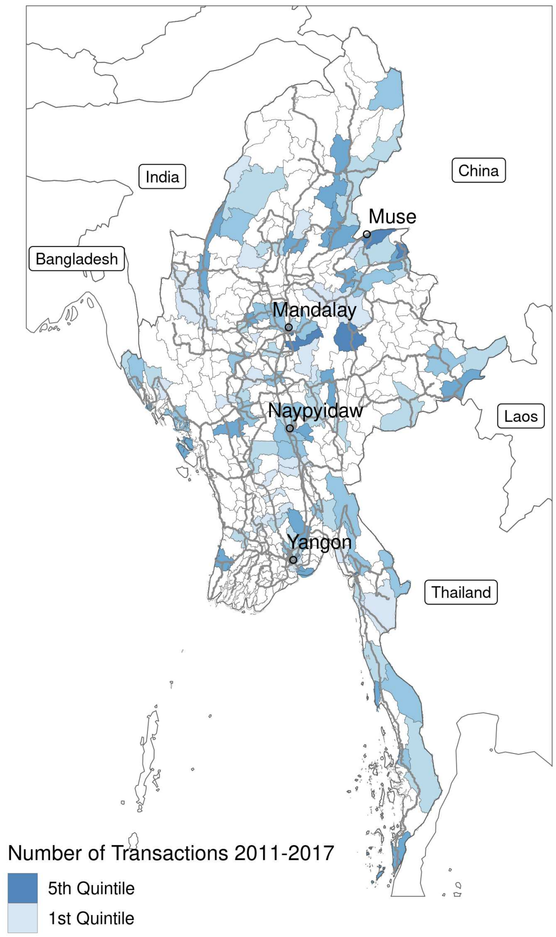


Fig. 3.6. Conflicts Blocking Trade Routes

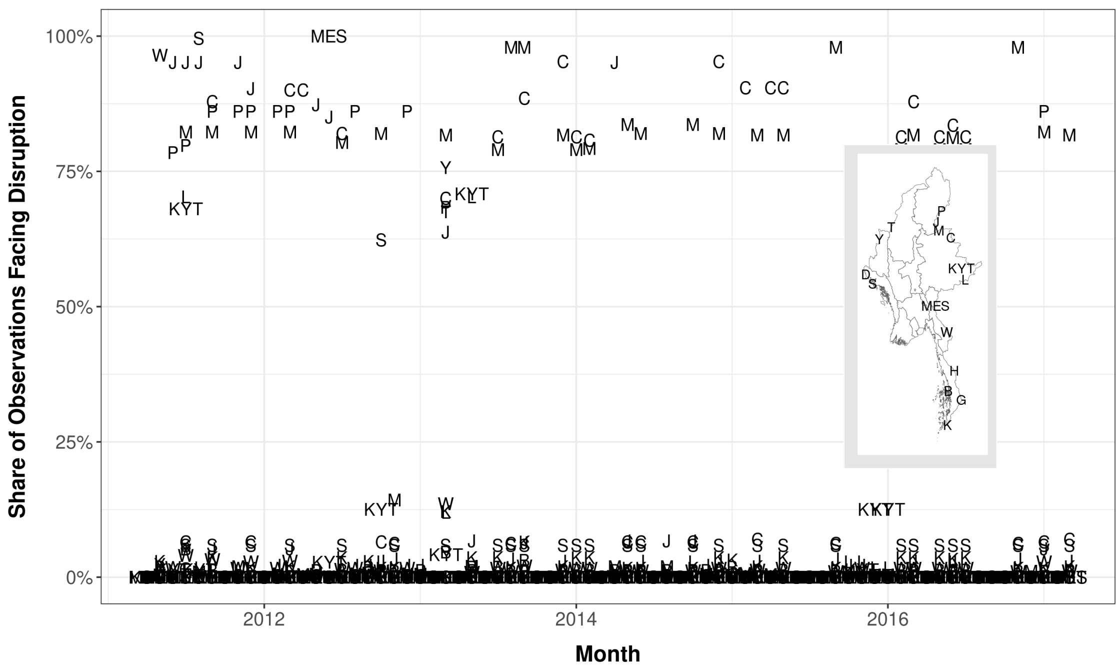


Fig. 3.7. Dynamic Specification

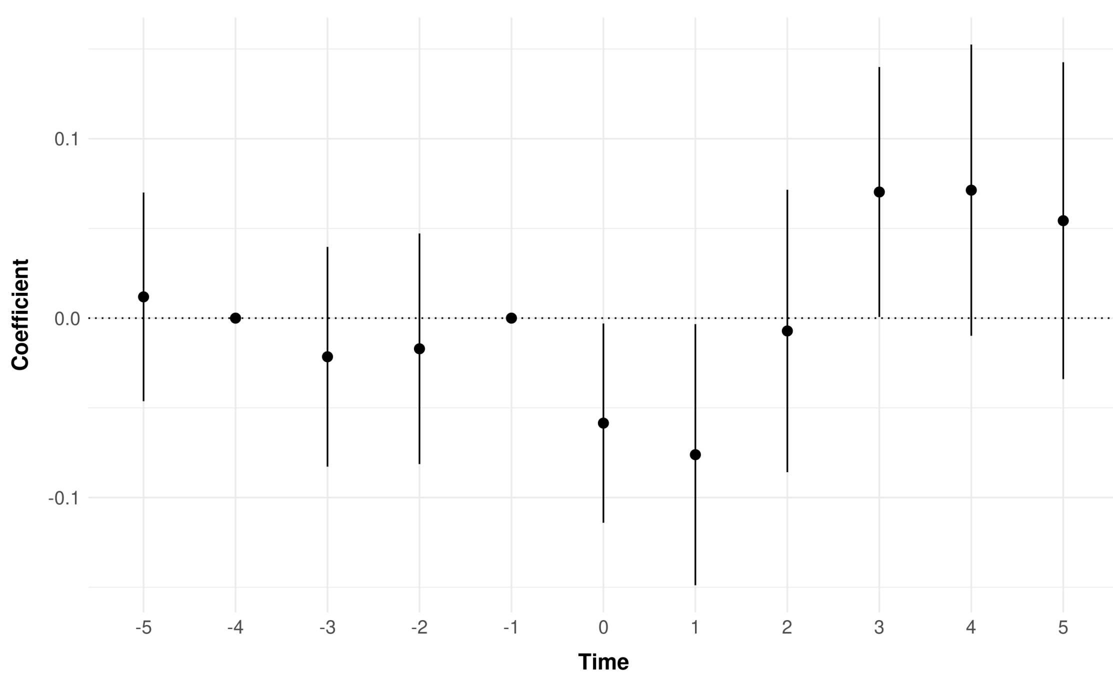
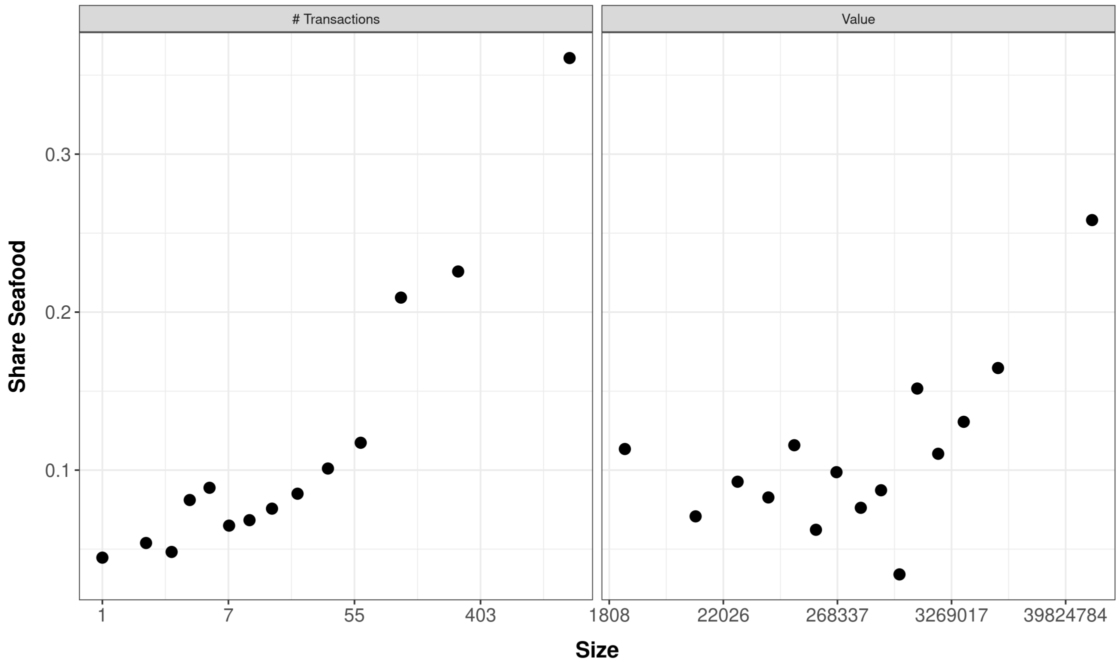


Fig. 3.8. Firm Size and Seafood Exports



C. Mapping Names to Ethnicities

The Burmese naming conventions are very unique, as they do not follow patronymic or matronymic conventions. Hence, there is no family name, only a given name. Burmese names also vary in the number of components/words. For example, ”Soe Mya” and ”Thin Thaw Tar Hnin” are both Burmese names. A further complication lies in the lack of convention on how to romanize Burmese names. Importantly for this classification exercise, names often contain honorific titles. These are often specific to just one ethnicity.

For classifying firm directors according to their ethnicity, I start with a sample of 8,660 unique director names. To classify names by ethnicity, I recruited multiple native Burmese research assistants (RAs) on *Upwork*, an online platform to hire freelancers. In a first round, I asked three RAs to classify a random sample of 1000 names by the eight major ethnicities in Myanmar: Chin, Shan, Kachin, Kayin, Kayah, Bamar, Rakhine and Mon. Some directors have foreign names. I thus also include a category ”Foreign”.

The majority of names in the dataset is Bamar. To identify and drop Bamar names in the wider dataset, I took names of this initial sample where all three RAs agreed that the name is Bamar. I then split these Bamar names into their components and identify names in the

larger sample where all components belong to these set of Bamar name components. I then drop these names identified as Bamar, which amounts to 58% of the sample.

I pass the remaining part of the sample to a set of three new RAs. If at least two RAs agree on the ethnicity, I keep this ethnicity. This leaves me with 97.3% of names being matched to an ethnicity or classified as foreign. As table 3.21 shows, 71% of firm directors are Bamar. The largest minority group is the Shan ethnicity making up 6% of the sample. There are no reliable estimates of the population share of the different ethnicities. While this information was recorded in the 2014 census, it was never released due to the political sensitivity of these numbers.[19] The CIA Factbook gives the following estimate: Bamar 68%, Shan 9%, Karen/Kayin 7%, Rakhine 4%, Chinese 3%, Indian 2%, Mon 2%, other 5%. The share of the Bamar majority is very similar in this population estimate and in the list of directors. Also, the Shan and the Karen minority appear as first and second largest ethnic minority in both. As expected, the Rakhine/Rohingya minority is vastly underrepresented in firm directors which is consistent with the discrimination they face in Burmese society.

Table 3.21: Number of Director Names by Ethnicity

Ethnicity	Count	Share
Bamar	6137	0.71
Foreign	1546	0.18
Shan	517	0.06
	233	0.03
Kayin	136	0.02
Kachin	54	0.01
Chin	32	0.00
Mon	4	0.00
Rakhine	1	0.00

D. Comparing Conflict Datasets

As mentioned in section 3, there are two main datasets capturing information on conflict at a high temporal and spatial resolution. In particular, *Uppsala Conflict Data Program* (UCDP) and the *Armed Conflict Location and Event Data Project* (ACLED) exist. The latter captures a wider variety of events. In particular, the former includes all violent events resulting in at least one death of state-based, non-state-based or one-sided violence conflicts which

[19]See Ferguson (2015) for a discussion on the controversy surrounding the 2014 census.

Fig. 3.9. Deaths in UCDP and ACLED datasets
The figure shows the (inverse hyperbolic sine of) cumulative deaths for Myanmar as recorded in the UCDP and ACLED datasets.

themselves lead to at least 25 battle deaths per year. ACLED imposes no such fatality requirement for the inclusion in the database. In Myanmar, this means that relatively minor skirmishes between actors that do not clash repeatedly are only recorded by ACLED. ACLED also records non-violent political events such as peaceful protests.

Overall, ACLED is thought of as being somewhat more extensive at the expense of containing more noise. Most relevant for my analysis, Eck (2012) argue that UCDP has higher geographic precision and should thus be preferred when doing subnational analysis. Below I show the differences between the two datasets in both the temporal and spatial dimension in the context of Myanmar.

To assess the correlation of both datasets, I aggregate the number of conflict fatalities to the monthly level. Figure 3.9 shows that the two datasets track each other closely.

Then, I divide Myanmar into a grid and count the number of fatalities occurring in each grid cell for both datasets over the same time period as above. I compute the correlation across grid cells, indicating the degree of spatial correlation among the datasets. Table

3.22 shows the correlation coefficient between the inverse hyperbolic sine of fatalities for different grid cell sizes. For relatively small grid cells, 0.25×0.25 degrees (roughly 25×25km) the correlation is 0.32. Even for large grid cells of 1×1 degree the correlation is only 0.71. This indicates substantial spatial differences between the two datasets. As depicted in the third column of the table, when focusing on locations with precisely coded location in both datasets (precision 1 or 2), the correlation coefficient increases especially for smaller grid cells, but remains below 0.75 for all levels of aggregation[20].

Table 3.22: UCDP vs ACLED - Spatial Correlation

Grid Cell Size (in Degrees)	Correlation	Correlation (precisely coded location)
0.25	0.32	0.45
0.33	0.39	0.49
0.50	0.54	0.59
1.00	0.71	0.74

[20]The correlation coefficients are similar when focusing only on the most precisely coded locations (precision 1).

Institut d’études politiques de Paris

ÉCOLE DOCTORALE DE SCIENCES PO

Programme doctoral en économie

Département d’économie

Doctorat en sciences économiques

Essais sur l’innovation et le commerce

Stefan PAULY

Thèse dirigée par

Thomas CHANEY, Professeur des Universités, Sciences Po

soutenue le 29 juin, 2023

Jury

Davide CANTONI (*rapporteur*)
Professor of Economics and Economic History, Ludwig-Maximilians-Universität München

Thomas CHANEY (*directeur de thèse*)
Professeur des Universités, Sciences Po

Isabelle MÉJEAN (*présidente*)
Professeure des Universités, Sciences Po

Claudia STEINWENDER (*rapportrice*)
Professor of Economics, Ludwig-Maximilians-Universität München

Résumé

À partir de la fin des années 1980, avec le début de la littérature sur la croissance endogène, les économistes ont reconnu le rôle important du progrès technologique pour une croissance économique soutenue. Dans un article important Romer (1990) modélise le progrès technologique en fonction d'un investissement délibéré en capital humain dans un bien non rival et partiellement excluable. La non-rivalité des idées, une caractéristique clé du modèle de Romer, implique que les nouvelles idées créent des externalités positives. Les spillovers de connaissances sont également reconnus depuis longtemps comme une externalité clé des agglomérations (Marshal, 1920). Pourtant, malgré leur importance perçue, les spillovers de connaissances sont difficiles à mesurer de manière empirique.

Jaffe et al. (1993) ont fait un pas important dans cette direction en introduisant l'utilisation de données sur les brevets dans la littérature économique. Ils utilisent les citations de brevets pour étudier la localisation géographique des spillovers de connaissances. S'appuyant sur ces premiers travaux et sur la littérature subséquente, les deux premiers chapitres de cette thèse utilisent les données sur les brevets pour suivre de près la manière dont les inventeurs collaborent et s'inspirent des idées des autres. En particulier, les données sur les brevets contiennent des informations précises sur la localisation des équipes d'inventeurs, les connaissances préexistantes qu'ils citent et sur lesquelles ils s'appuient, et les entreprises qui déposent le brevet.

Le premier chapitre de cette thèse, co-écrit avec Fernando Stipanicic, étudie comment une augmentation de la mobilité humaine peut conduire à une plus grande diffusion des connaissances entre les différentes localités. Nous étudions ensuite comment ces transferts de connaissances se traduisent par la création de nouvelles idées. Cela permet d'identifier les paramètres clés de la fonction de production des idées que les premiers travaux théoriques n'ont pas été en mesure de déterminer en raison d'un manque de données détaillées. En outre, nous montrons que les villes initialement moins innovantes

sont celles qui bénéficient le plus d'une augmentation des spillovers de connaissances, ce qui entraîne une convergence.

Plutôt que d'étudier les spillovers de connaissances et ses effets au niveau géographique, une littérature récente décrit la production de la connaissance au sein de l'entreprise. Avec l'essor des entreprises multinationales, la production est de plus en plus dispersée entre les pays au sein d'une même entreprise, et ces entreprises sont responsables d'une grande partie de la création de connaissance. Le deuxième chapitre, qui est co-écrit avec Çagatay Bircan et Beata Javorcik, étudie comment les frictions de communication peuvent affecter la collaboration entre les inventeurs situés dans une filiale étrangère et ceux situés au siège de l'entreprise. Il s'agit d'un point crucial, car une grande partie des connaissances des entreprises est détenue par les inventeurs basés au siège et, comme nous le montrons, les inventeurs des filiales étrangères déposent des brevets de meilleure qualité s'ils collaborent avec le siège. Alors que le premier chapitre traite des transferts de connaissances entre entreprises, le deuxième chapitre s'intéresse aux déterminants de la diffusion des connaissances à l'intérieur de l'entreprise.

Comparé à la diffusion des connaissances et à l'échange d'idées, l'échange de biens est depuis longtemps reconnu comme un moteur essentiel de la croissance économique moderne. Une multitude d'articles étudie l'interaction entre les frictions bilatérales au commerce, l'échange de biens et ses effets sur le bien-être et la croissance. Alors qu'une grande partie de cette littérature se concentre sur les frictions commerciales de longue durée, telles que les tarifs ou la géographie, peu d'études se concentrent sur la manière dont les interruptions fréquentes et imprévisibles de l'accès au marché affectent la façon dont les entreprises échangent. Cette question est particulièrement pertinente dans les pays en développement où l'infrastructure peut être peu fiable et où le commerce est souvent dominé par des entreprises informelles et peu résilientes. Dans mon troisième chapitre, j'étudie la manière dont les entreprises du Myanmar gèrent la perte répétée d'accès aux marchés des pays voisins, causée par des conflits ethniques le long des routes principales. Je constate que les entreprises compensent partiellement les pertes en vendant sur le marché intérieur et en retardant les exportations. La taille de l'entreprise et les liens avec l'armée rendent les entreprises plus résistantes.

Chapitre 1: La création et la diffusion des connaissances: L'expérience de l'ère du jet

Dans le premier chapitre, co-écrit avec Fernando Stipanicic, nous examinons comment la mobilité des personnes influe sur la diffusion des idées. Nous étudions une avancée technologique majeure qui a facilité les déplacements de personnes: Dans les années 1950, le moteur à réaction a été introduit dans l'aviation civile, ce qui a permis de créer des avions plus rapides et avec un rayon d'action plus long. À l'époque, le réseau aérien était fortement réglementé, ce qui contraignait les compagnies aériennes à réoptimiser leurs itinéraires. Cela nous permet d'affirmer que les améliorations technologiques ont entraîné un choc exogène plausible sur le réseau existant.

Pour étudier cette période historique, nous numérisons les horaires de vol des principales compagnies aériennes de l'époque. Cela nous permet de calculer le temps de trajet exact par avion, avant et après l'arrivée des avions à réaction. Nous constatons que la durée des trajets entre les villes principales américaines a diminué d'un tiers entre 1951 et 1966. Nous combinons ces données avec celles relatives aux brevets attribués aux entreprises et à la localisation des inventeurs. Cela nous donne une image complète des brevets déposés par une entreprise, y compris le lieu où les connaissances ont été produites et les idées antérieures qui ont été citées.

Nous constatons que la forte augmentation de la mobilité s'est traduite par une circulation accrue des idées, en particulier sur les longues distances géographiques. Les citations de brevets entre établissements distants de plus de 2000 km augmentent de 8% en raison de la réduction du temps de déplacement, ce qui correspond à 38% de l'augmentation globale pour cet intervalle de distance.

L'augmentation de la diffusion des connaissances a également conduit à la création de nouvelles connaissances. En utilisant l'élasticité de la diffusion des connaissances par rapport au temps de déplacement estimée dans la première étape, nous construisons une mesure de l'accès aux connaissances. Nous montrons que le nombre de brevets augmente quand un secteur dans un lieu particulier accède à des nouvelles connaissances grâce à la réduction des temps de déplacement. En outre, les résultats montrent que cet effet est plus important pour les sites initialement moins innovants, ce qui implique une convergence entre les sites.

En résumé, cet épisode historique nous permet d'établir un lien de causalité entre la durée des trajets et la diffusion des connaissances. L'abaissement des frictions de mobilité augmente la circulation des idées. Nous montrons qu'une diffusion accrue des connaissances est cruciale pour les nouvelles innovations, en particulier pour les inventeurs situés en dehors des grands centres d'innovation.

Chapitre 2: Time after Time: les coûts de communication et la collaboration entre inventeurs au sein de l'entreprise multinationale

Dans le deuxième chapitre, qui est un travail conjoint avec Çağatay Bircan et Beata Javorcik, nous examinons la production de connaissances au sein des entreprises multinationales. Pour étudier les activités globales de recherche et développement des entreprises, nous combinons les données sur les brevets avec les liens de propriété des entreprises, ce qui nous permet d'attribuer chaque brevet à son propriétaire final et à son lieu de production grâce aux informations sur le lieu de résidence des inventeurs.

Nous constatons tout d'abord qu'une grande partie des activités de R&D des entreprises se déroulent en dehors du pays du siège. Par exemple, seulement 61% des brevets les plus précieux déposés par des entreprises américaines sont produits par des équipes d'inventeurs exclusivement basées aux États-Unis. En nous concentrant sur les brevets impliquant une filiale étrangère, nous montrons que les inventions sont de meilleure qualité si elles sont créées en collaboration avec le siège. Cela suggère que l'accès aux connaissances du siège peut rendre les inventeurs basés dans les filiales plus productifs, et cela souligne l'importance des frictions dans cette diffusion des connaissances.

Nous constatons que des différences de fuseau horaire plus importantes entre une filiale et le siège rendent les collaborations moins probables. Un doublement du chevauchement des heures d'ouverture augmente la probabilité de collaboration de 5 points de pourcentage, et ce d'autant plus dans les catégories technologiques expérimentales qui nécessitent probablement une fréquence élevée d'interactions. L'effet reste relativement stable après l'inclusion de la distance physique comme variable de contrôle. L'ensemble de ces résultats suggère que les différences de fuseaux horaires constituent une entrave à la communication, distincte des frais de déplacement.

Motivés par l'augmentation de la part des brevets en collaboration transfrontalière et pour fournir des preuves causales du rôle des frictions de communication, nous étudions les épisodes de libéralisation du secteur des télécommunications dans les années 1980 et 1990. Cette déréglementation a entraîné une baisse significative du prix des appels internationaux. En rassemblant des données sur les dates de libéralisation et les prix des appels bilatéraux, nous montrons que les collaborations entre les filiales étrangères et les sièges sociaux augmentent lorsque les marchés des télécommunications sont libéralisés dans les deux pays et que les prix des appels chutent. Nous montrons ensuite que l'effet des appels internationaux moins chers se concentre sur les paires de sites où les heures d'ouverture sont similaires et sur les technologies expérimentales, ce qui est cohérent avec nos résultats transversaux. Enfin, nous montrons que l'effet est concentré sur les inventeurs moins expérimentés.

En résumé, nous montrons que les différences de fuseaux horaires constituent un obstacle à la diffusion des connaissances au sein des entreprises multinationales. Si l'amélioration des technologies de communication peut faciliter la collaboration transfrontalière des inventeurs, les fuseaux horaires continueront à jouer un rôle important. Nous montrons que cela est particulièrement vrai pour les technologies qui nécessitent une fréquence de communication élevée et pour les inventeurs moins expérimentés.

Chapitre 3 : Des commerçants en difficulté sur la route de Mandalay

Enfin, dans un troisième article, j'étudie la manière dont l'incertitude vis-à-vis des frictions commerciales affecte l'échange de biens. Je combine des données commerciales au niveau des transactions en Birmanie avec des données sur les conflits. La Birmanie a une longue histoire de conflits ethniques qui se concentrent dans ses régions frontalières. Mais il dépend fortement des exportations vers ses voisins (la Chine et la Thaïlande, en particulier), dont une partie emprunte la voie terrestre et traverse donc les zones touchées par le conflit. En traitant l'émergence d'insurrections dans les régions frontalières comme des chocs soudains et inattendus sur les coûts commerciaux, je confirme les preuves anecdotiques que les conflits affectent négativement le commerce de transit.

Ensuite, je montre que les exportations des entreprises transitent généralement par une seule ville frontalière. Ce manque de diversification rend les entreprises vulnérables aux perturbations causées par les conflits le long d'une seule route. En conséquence, les prix

locaux des produits d'exportation chutent lorsque des routes importantes sont perturbées, ce qui suggère que les entreprises cessent d'exporter et vendent plutôt une partie de leur produits sur les marchés locaux. Le commerce diminue davantage pour les produits périssables qui sont plus risqués à exporter lorsque la durée du trajet jusqu'à la frontière est incertaine. Pour les produits durables, j'observe que les entreprises regroupent leurs expéditions et augmentent la taille moyenne de leurs envois, ce qui les protège des pertes.

En utilisant la périssabilité des produits pour identifier la manière dont les différentes entreprises opèrent dans l'incertitude, je constate que les grandes entreprises sont plus susceptibles de continuer à exporter. En outre, les liens avec l'armée, un acteur puissant en Birmanie, sont particulièrement précieux en période de conflit et pour les entreprises qui commercialisent des produits périssables.

En résumé, ce projet montre que la combinaison d'un manque de diversification des marchés d'exportation et de perturbations fréquentes pèse lourdement sur les gains à l'exportation, affectant les entreprises et les secteurs de manière hétérogène. Alors que la fin des conflits ethniques en Birmanie n'est pas en vue, la réduction des barrières à l'entrée vers de nouvelles routes d'exportation et un marché intérieur fort augmenteraient de manière significative la résilience des entreprises.

References

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Marshal, A. (1920). Principles of economics. *London: Mcmillan*.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102.